



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 96045

TO: Mark Navarro  
Location: 8A15  
Art Unit: 1645  
Friday, June 06, 2003

Case Serial Number: 09/150947

From: Barb O'Bryen  
Location: Biotech-Chem Library  
CM1-6A05  
Phone: 308-4291 *BOB*

barbara.obryen@uspto.gov

### Search Notes

Mark,  
all of the requested seqs contain bunches of Xaa's. With standard search parameters, Xaa's in either the query seq or the database seq are considered non-matches, so I don't know how useful you will find these results to be. We have an alternative search that can be done in cases like this, called "DX matching table". With this matching table, Xaa in the query sequence matches anything in the database seq. Xaa in a database seq still is considered a non-match.  
If you think DX matching table results would be more useful to you, please let me know.  
Barb

#### O'Bryen, Barbara

From:  
Sent:  
To:  
Subject:

Navarro, Albert  
Thursday, June 05, 2003 1:38 PM  
O'Bryen, Barbara  
09/150,947

Mark Navarro  
1645  
306-3225  
8A15

Hello Barb,

Could I trouble you for a search and an interference search of:

SEQ ID NO: 13-18

Thanks

Mark



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:47:43 ; Search time 33.9394 Seconds

(without alignments)  
39.261 Million cell updates/sec

Title: US-09-150-947f-13

Sequence: 1 KXXXTXQEXD 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

1: A.Geneseq\_101002.\*  
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	24	77.4	233	21	AAV93741	Amino acid sequence
2	24	77.4	258	23	ABP29565	Streptococcus poly
3	23	74.2	17	23	ABP76254	Staphylococcus aur
4	23	74.2	24	21	AAV97855	Staphylococcus aur
5	23	74.2	24	22	AAV72193	Peptide from regio
6	23	74.2	29	22	AAV59147	Human brain expro
7	23	74.2	29	22	AAV71683	Human bone marrow
8	23	74.2	29	22	AAV31976	Peptide #6013 enco
9	23	74.2	29	23	ABG41496	Human peptide enco
10	23	74.2	228	22	AAV63856	Amino acid sequence

11	23	74.2	238	12	AAV13208	Staphylococcal ent
12	23	74.2	238	14	AAV45016	Staphylococcal ent
13	23	74.2	238	22	AAV67343	Staphylococcus aur
14	23	74.2	238	23	ABP76239	Staphylococcus aur
15	23	74.2	239	12	AAV13207	Staphylococcal ent
16	23	74.2	239	14	AAV45015	Staphylococcal ent
17	23	74.2	239	20	AAV06254	Staphylococcal gro
18	23	74.2	239	20	AAV06255	Staphylococcal gro
19	23	74.2	239	20	AAV06256	Staphylococcal gro
20	23	74.2	239	20	AAV06257	Staphylococcal gro
21	23	74.2	239	20	AAV06258	Staphylococcal gro
22	23	74.2	239	20	AAV06251	Staphylococcal gro
23	23	74.2	239	20	AAV06252	Staphylococcal gro
24	23	74.2	239	20	AAV06253	Staphylococcal gro
25	23	74.2	239	22	AAV67342	Staphylococcus aur
26	23	74.2	239	23	ABP76238	Staphylococcus aur
27	23	74.2	266	21	AAV70108	Staphylococcal ent
28	23	74.2	266	23	ABP79507	Staphylococcal SEC
29	23	74.2	311	22	ABG17229	Novel human diagno
30	23	74.2	405	21	AAV26155	B. cereus zwitterm
31	22	71.0	10	19	AAV64639	Synthetic SEB pept
32	22	71.0	10	19	AAV64638	Synthetic SEB pept
33	22	71.0	12	19	AAV64636	Synthetic SEB pept
34	22	71.0	12	19	AAV64637	Synthetic SEB pept
35	22	71.0	13	19	AAV64640	Synthetic SEB pept
36	22	71.0	13	19	AAV64641	Synthetic SEB pept
37	22	71.0	13	19	AAV64646	Synthetic SEB pept
38	22	71.0	14	19	AAV64644	Synthetic SEB pept
39	22	71.0	14	19	AAV64645	Synthetic SEB pept
40	22	71.0	17	23	ABP76253	Staphylococcus aur
41	22	71.0	17	23	ABP76257	Streptococcus pyog
42	22	71.0	17	23	ABP76258	Enterotoxin, pyoge
43	22	71.0	24	19	AAV64642	Synthetic SEB pept
44	22	71.0	24	21	AAV97854	Staphylococcus aur
45	22	71.0	24	21	AAV97859	Group A streptococ

#### ALIGNMENTS

RESULT 1  
ID AAV93741 standard; Protein: 233 AA.  
AC AAV93741:  
XX  
XX 03-OCT-2000 (first entry)  
XX  
XX Amino acid sequence of the mature SMEZ-2 superantigen protein.  
DE  
XX Sperantigen; SMEZ-2; SPE-G; SPE-H; SPE-J; Streptococcal disease;  
XX Kwasaki syndrome; T cell activation; cancer therapy.  
KM  
XX Streptococcus pyogenes.  
OS  
XX  
XX WO200039159-A1.  
PN  
XX 06-JUL-2000.  
PD  
XX 24-DEC-1999; 99WO-N00228.  
PF  
XX 24-DEC-1998; 98NZ-0333589.  
PR  
XX (AUCK-) AUCKLAND UNISERVICES LTD.  
PA  
XX Fraser JD, Proft T;  
XX WPI, 2000-452370/39.  
PI  
XX N-PSDB; AAA47147.  
DR  
XX Novel superantigens from streptococcus pyogenes useful for genotyping  
XX streptococcus pyogenes clones expressing SMEZ-2 and for diagnosing a  
PT Kawasaki syndrome  
PT

XX Claim 2; Fig 2; 72pp; English.  
 PS  
 CC The present sequence represents the SMEZ-2 superantigen protein. The  
 CC specification describes superantigen proteins SMEZ-2, SPE-G, SPE-H  
 CC and SPE-I. The superantigen polynucleotides and polypeptides are  
 CC used for subtyping Streptococci. They are also used for diagnosing  
 CC Streptococcal disease. The superantigens are used in diagnosis of  
 CC disease such as Kawasaki syndrome. They are also useful to recruit  
 CC and activate T cells in a relatively non-specific fashion since  
 CC they bind a large number of T cell receptor molecules by binding to the  
 CC Vbeta domain. Superantigen constructs are useful in cancer therapy.  
 CC  
 SQ Sequence 233 AA;  
 QY Query Match 77.4%; Score 24; DB 21; Length 233;  
 Best Local Similarity 50.0%; Pred. No. 73;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 DB 1 KXXXTXQEXD 10  
 149 KTTVTAQED 158  
 RESULT 2  
 ID ABB29565 standard; Protein; 258 AA.  
 AC ABB29565;  
 DT 02-JUL-2002 (first entry)  
 DE Streptococcus polypeptide SEQ ID NO 8306.  
 XX  
 KW Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;  
 KW group A streptococcus; Streptococcus pyogenes; antihemolytic;  
 KM antiinflammatory; infection; vaccine; meningitis; gene therapy.  
 XX  
 OS Streptococcus pyogenes.  
 XX  
 PN WO200234771-A2.  
 PD 02-MAY-2002.  
 PF 29-OCT-2001; 2001WO-GB04789.  
 PR 27-OCT-2000; 2000GB-0026333.  
 PR 24-NOV-2000; 2000GB-0028727.  
 PR 07-MAR-2001; 2001GB-0005640.  
 PA (CHIR-) CHIRON SPA.  
 PA (GENO-) INST GENOMIC RES.  
 PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;  
 PI Tettelin H;  
 DR WPI: 2002-352536/38.  
 DR N-PSDB; ABB70196.  
 XX  
 PT New Streptococcus protein for the treatment or prevention of infection  
 PT or disease caused by Streptococcus bacteria, such as meningitis, and  
 PT for detecting a compound that binds to the protein -  
 PS  
 PS Claim 1; Page 3947; 4525pp; English.  
 CC The invention relates to a protein (ABP25413-ABP30895) from group B  
 CC streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS  
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
 CC the specification. The proteins have antibacterial and antiinflammatory  
 CC activity. (1), nucleic acids encoding (1), ABB66044-ABB71526 and  
 CC antibodies that bind (1) are used in the manufacture of medicaments for  
 CC the treatment or prevention of infection or disease caused by  
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

CC Nucleic acids encoding (1) are used to detect Streptococcus in a  
 CC biological sample. (1) is used to determine whether a compound binds to  
 CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be  
 CC used as a vaccine or diagnostic composition. The disease caused by  
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
 CC acid encoding (1) may be used to recombinantly produce (1) and may be  
 CC used in gene therapy. Antibodies to (1) are used for affinity  
 CC chromatography, immunoassays, and distinguishing/identifying  
 CC Streptococcus proteins.  
 CC  
 SQ Sequence 258 AA;  
 QY Query Match 77.4%; Score 24; DB 23; Length 258;  
 Best Local Similarity 50.0%; Pred. No. 81;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 DB 1 KXXXTXQEXD 10  
 174 KTTVTAQED 183  
 RESULT 3  
 ID ABB76254 standard; Peptide; 17 AA.  
 AC ABB76254;  
 DT 09-AUG-2002 (first entry).  
 DE Staphylococcus aureus enterotoxin C1 peptide fragment.  
 XX  
 KW Enterotoxin C; SECI; superantigen; antigen; tumour; cancer;  
 KW antitumour; therapy.  
 XX  
 OS Staphylococcus aureus.  
 XX  
 PN US2002051765-A1.  
 PD 02-MAY-2002.  
 PF 19-DEC-2000; 2000US-0741503.  
 PR 31-JAN-1994; 94US-0189424.  
 PR 19-JUN-1995; 95US-0491746.  
 PR 03-OCT-1989; 89US-0416530.  
 PR 17-JAN-1990; 90US-0465577.  
 PR 17-JAN-1991; 91WO-US00342.  
 PR 01-JUN-1992; 92US-0891718.  
 PR 02-MAR-1993; 93US-0025144.  
 PA (TERM/) TERMAN D S.  
 PA  
 PI Terman DS;  
 DR WPI: 2002-415196/44.  
 XX  
 PT Reagent for treating cancer without the need for e.g. radiotherapy,  
 PT comprises a specific V beta subset of T cells sensitized to a growing  
 PT tumor and stimulated with superantigens -  
 PS  
 PS Disclosure; Page 4; 17pp; English.  
 CC The present sequence corresponds to amino acids 151-167 of  
 CC enterotoxin C1 (SECI, see ABB76238) of Staphylococcus aureus.  
 CC Sequence similarities are found in the corresponding regions  
 CC of other staphylococcal enterotoxins and streptococcal pyrogenic  
 CC exotoxins (see ABB76252-59). In the present invention, synthetic  
 CC polypeptides useful in tumour therapy and in blocking or destroying  
 CC autoreactive T and B lymphocyte populations are characterised by  
 CC substantial structural homology to staphylococcal enterotoxin A and  
 CC enterotoxin B, and to streptococcal pyrogenic exotoxins, with  
 CC statistically significant sequence homology and similarity,  
 CC including alignment of cysteine residues and similar hydrophaty



CC profiles. These superantigens are used to treat solid tumours,  
 CC including their metastases, without radiation, surgery or standard  
 CC chemotherapeutic agents. A claimed method of human cancer treatment  
 CC involves contacting haematopoietic cells from a patient with one or  
 CC more superantigens ex vivo to generate stimulated cells, selecting  
 CC a specific V beta subset of cells, and reintroducing these cells  
 CC into the patient to induce an in vivo therapeutic, tumouricidal  
 CC reaction.

XX Sequence 17 AA:

Query Match 74.2%; Score 23; DB 23; Length 17;  
 Best Local Similarity 50.0%; Pred. No. 9.8;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTQEXD 10  
 Db 1 KKSVAQELD 10

#### RESULT 4

ID AAY97855 standard; peptide: 24 AA.

XX AAY97855;

DT 29-AUG-2000 (first entry)

DE Staphylococcus aureus enterotoxin SEC, conserved region 2.

XX Staphylococcal infection; enterotoxin; septic shock; toxic shock;

KW vaccine; antibody.

XX Staphylococcus aureus.

XX W0200020598-A1.

PD 13-APR-2000.

PE 24-SEP-1999; 99WO-US22180.

PR 07-OCT-1998; 98US-0168303.

PR 18-JUN-1999; 99US-0335381.

PA (UYRQ ) UNIV ROCKEFELLER.

PI Bannan JD, Visvanathan K, Zabriskie JB;

DR WPI: 2000-303782/26.

PT Peptides useful for preventing and reducing the symptoms of toxic shock

PS infections -

XX Example 1; Page 72; 115pp; English.

XX The invention relates to novel peptides (AAY97838-Y97843) comprising a

CC consensus amino acid sequence derived from two conserved regions

CC (regions 1 and 2) of Staphylococcus aureus enterotoxins and

CC streptococcal pyrogenic toxins. Consensus region 1a (a preferred

CC (AAY97863) and consensus region 2a (a preferred consensus region 2) has

CC the sequence K-X6-X7-X8-

CC X9-X10-X11-X12-X13-D-X14-X15-X16-R-X17-X18-X27-X19-X20-X21-X22-X23-X24-Y

CC (AAY97864), where: X1, X8, X13 and X24 are each independently selected

CC from L, I and V; X2, X4, X5, X6, X7, X9, X10, X11, X12, X14, X15, X16,

CC X17, X18, X19, X20, X21, X22 and X23 are each independently selected from

CC the group consisting of any amino acid;

CC X3, X5 and X26 are each independently selected from the group consisting

CC of any amino acid and of no amino acid;

CC X27 is either L or Y.

CC The invention also relates to serum antibodies induced by the peptides

CC which provide protection against, or reduce the severity of toxic shock

CC and septic shock caused by the staphylococcal and streptococcal  
 CC pyrogenic toxins. The pyrogenic exotoxins of Group A streptococci and  
 CC the enterotoxins of Staphylococcus aureus (which are also pyrogenic  
 CC exotoxins) constitute a family of structurally related toxins which  
 CC share similar biological activities. They stimulate CD4+, CD8+ and  
 CC gamma-delta+ T-cells by binding the beta-chain variable region (V-beta)  
 CC elements on the lateral face of the T-cell receptor (TCR) while  
 CC simultaneously binding the lateral face of the class II major  
 CC histocompatibility complex (MHC) of antigen presenting cells. This causes  
 CC aberrant proliferation of T-cells, which stimulates other components of  
 CC the immune system, causing injury to the host. The peptides are used to  
 CC prevent, treat or protect against toxic shock and septic shock resulting  
 CC from bacterial infections in mammals, particularly humans. The peptides  
 CC are used for inducing serum antibodies that bind at least one  
 CC staphylococcal enterotoxin or streptococcal exotoxin and both the  
 CC peptides and antibodies can be used in diagnostic assays to aid in the  
 CC diagnosis of disease related to the presence of bacterial toxins. Nucleic  
 CC acids encoding a peptide of the invention can be used for the production  
 CC of the peptides for diagnostic reagents, as vaccines and for therapies  
 CC for pyrogenic exotoxin related diseases. Vectors expressing high levels  
 CC of the peptides can be used in immunotherapy and immunoprophylaxis when  
 CC expressed in humans. The antibodies are used for passive immunisation  
 CC therapy to prevent or increase resistance to toxic shock syndrome or  
 CC septic shock and to ameliorate the effects of conditions associated with  
 CC the presence of staphylococcal or streptococcal pyrogenic toxins. The  
 CC amino acid sequences of the peptides are sufficiently common that they  
 CC can be used for eliciting antibodies which are cross-reactive with toxins  
 CC derived from various bacteria. Sequences AAY97853-Y97861 represent  
 CC conserved region 2 of various Staphylococcus aureus enterotoxins and  
 CC Group A streptococcal pyrogenic exotoxins.

XX Sequence 24 AA:

Query Match 74.2%; Score 23; DB 21; Length 24;

Best Local Similarity 50.0%; Pred. No. 14;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTQEXD 10  
 Db 1 KKSVAQELD 10

#### RESULT 5

ID AAY72193 standard; peptide: 24 AA.

XX AAY72193;

DT 24-APR-2001 (first entry)

DE Peptide from region 2 of staphylococcal enterotoxin SEC.

XX Mononuclear cell; blastogenesis; inhibitor; HIV; replication; therapy;

XX staphylococcal enterotoxin; human immunodeficiency virus; T cell;

XX autoimmune disease; immunisation.

XX Staphylococcus sp.

XX W0200078790-A2.

PD 28-DEC-2000.

PE 16-JUN-2000; 2000WO-US16680.

PR 18-JUN-1999; 99US-0336627.

PA (UYRQ ) UNIV ROCKEFELLER.

PI Visvanathan K, Zabriskie JB;

DR WPI: 2001-080820/09.

PT Providing protection against, and reducing the severity of, human

PT Immunodeficiency virus infections and associated deleterious effects,  
PT using peptides from homologous sequences of staphylococcal and  
XX streptococcal toxins  
PS Disclosure; Page 36; 76pp; English.  
XX  
CC The present sequence is a peptide from region 2 of staphylococcal  
CC enterotoxin SEC.  
CC The peptide, nucleic acid encoding the peptide and antibody (Ab) produced  
CC against the peptide are useful for inhibiting biogenesis of  
CC mononuclear cells (eg. T cells) in the presence of human immunodeficiency  
CC virus (HIV), inhibiting HIV replication and protecting a mammal against  
CC the deleterious effects of HIV. The peptide is also used to ameliorate  
CC the effects of autoimmune diseases associated with the presence of HIV.  
CC The Ab is used for passively immunizing a mammal against the deleterious  
CC effects of HIV.  
XX  
SQ Sequence 24 AA;  
Query Match 74.2%; Score 23; DB 22; Length 24;  
Best Local Similarity 50.0%; Pred. No. 14;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
OY 1 KXXXTXQEXD 10  
DB 1 KKSVTAEQED 10  
RESULT 6  
ID AAMS9147  
XX AAMS9147 standard; Protein; 29 AA.  
XX  
AC AAMS9147;  
XX  
DT 05-NOV-2001 (first entry)  
XX  
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 31252.  
XX  
XX Human; brain expressed exon; gene expression analysis; probe;  
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
XX epilepsy; cancer.  
OS Homo sapiens.  
XX  
PN WO200157275-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00667.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI: 2001-483446/52.  
XX  
PT Single exon nucleic acid probes for analyzing gene expression in human  
PT brains -  
XX  
PS Example 4; SEQ ID NO: 31252; 650pp + Sequence Listing; English.  
XX  
CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is a protein encoded by one of  
CC the probes of the invention.  
XX  
SQ Sequence 29 AA;  
Query Match 74.2%; Score 23; DB 22; Length 29;  
Best Local Similarity 50.0%; Pred. No. 17;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
OY 1 KXXXTXQEXD 10  
DB 10 KALSTGQED 19  
RESULT 7  
ID AAM71683  
XX AAM71683 standard; Protein; 29 AA.  
XX  
AC AAM71683;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 31989.  
XX  
XX Human; bone marrow expressed exon; gene expression analysis; probe;  
KW microarray; cancer; leukaemia; lymphoma; myeloma.  
XX  
OS Homo sapiens.  
XX  
PN WO200157276-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00668.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI: 2001-488900/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human bone marrow -  
XX  
PS Example 4; SEQ ID NO: 31989; 650pp + Sequence Listing; English.  
XX  
CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC bone marrow. They can be used to measure gene expression in bone marrow  
CC samples, which may enable the improved diagnosis and treatment of cancers  
CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
CC protein encoded by one of the probes of the invention.  
XX  
SQ Sequence 29 AA;  
Query Match 74.2%; Score 23; DB 22; Length 29;  
Best Local Similarity 50.0%; Pred. No. 17;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
OY 1 KXXXTXQEXD 10  
DB 10 KALSTGQED 19

RESULT 8  
 AAM31976  
 ID AAM31976 standard; Protein: 29 AA.  
 AC AAM31976;  
 XX  
 XX 17-OCT-2001 (first entry)  
 DE Peptide #6013 encoded by probe for measuring placental gene expression.  
 XX  
 XX Probe: microarray; human; placenta; antenatal diagnosis;  
 KM genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157272-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00663.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PT Penn SG, Hanzel DK, Chen W, Rank DR;  
 DR WPI; 2001-488897/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 XX analyzing gene expression in human placenta -  
 PS Claim 27; SEQ ID NO 32245; 654pp; English.  
 CC The present invention relates to single exon nucleic acid probes (SENP;  
 CC see AAI31315-AA157546). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for producing a microarray for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from human placenta. The probes are useful for antenatal diagnosis of  
 CC human genetic disorders.  
 CC  
 SO Sequence 29 AA:  
 Query Match 74.2%; Score 23; DB 22; Length 29;  
 Best Local Similarity 50.0%; Pred. No. 17;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 KXXXXXQEND 10  
 DB 10 KALSTQGEID 19  
 RESULT 9  
 ABG41496  
 ID ABG41496 standard; Peptide: 29 AA.  
 AC ABG41496;  
 XX  
 XX 19-AUG-2002 (first entry)  
 DE Human peptide encoded by genome-derived single exon probe SEQ ID 31161.  
 XX  
 XX Human; single exon probe; asthma; lung cancer; COPD; ILD;  
 KM chronic obstructive pulmonary disease; interstitial lung disease;  
 KM familial idiopathic pulmonary fibrosis; neurofibromatosis;  
 KM tuberos sclerosis; Gaucher's disease; Niemann-Pick disease;  
 KM Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemostiderosis;

KM pulmonary histiocytosis; lymphangioleiomyomatosis; Kargener syndrome;  
 KM pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
 KM primary ciliary dyskinesia; pulmonary hypertension;  
 XX hyaline membrane disease.  
 OS Homo sapiens.  
 XX  
 PN WO200186003-A2.  
 XX  
 PD 15-NOV-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00665.  
 XX  
 PR 04-FEB-2000; 2000US-180312P.  
 PR 26-MAY-2000; 2000US-207456P.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-234687P.  
 PR 27-SEP-2000; 2000US-236359P.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PT Penn SG, Hanzel DK, Chen W, Rank DR;  
 DR WPI; 2002-114183/15.  
 XX  
 PT Spatially-addressable set of single exon nucleic acid probes, used to  
 XX measure gene expression in human lung samples -  
 PS Claim 27; SEQ ID NO 31161; 634pp; English.  
 XX  
 CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human lung comprising single exon nucleic acid probes having one of  
 CC 12614 nucleic acid sequences mentioned in the specification, or their  
 CC complements or the 12387 open reading frames derived from the 12614  
 CC probes. Also included are a microarray comprising the novel set of  
 CC probes; the novel set of probes which hybridize at high stringency to a  
 CC nucleic acid expressed in the human lung; measuring gene expression in a  
 CC sample derived from human lung, comprising (a) contacting the array with  
 CC a collection of detectably labeled nucleic acids derived from human lung  
 CC mRNA, and (b) measuring the label detectably bound to each probe of  
 CC the array; identifying exons in a eukaryotic genome, comprising  
 CC (a) algorithmically predicting at least one exon from genomic sequences  
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably  
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
 CC having a fragment identical to the predicted exon, the probe is included  
 CC in the above mentioned microarray; assigning exons to a single gene,  
 CC comprising (a) identifying exons from genomic sequence by the method  
 CC above and (b) measuring the expression of each of the exons in several  
 CC tissues and/or cell types using hybridisation to a single exon  
 CC microarray having a probe with the exon, where a common pattern of  
 CC expression of the exons in the tissues and/or cell types indicates that  
 CC the exons should be assigned to a single gene; a peptide comprising one  
 CC of 12011 sequences, mentioned in the specification, or encoded by the  
 CC probes/open reading frames (ORF). The probes are used for gene  
 CC expression analysis, and for identifying exons in a gene, particularly  
 CC using human lung derived mRNA and for the study of lung diseases  
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease  
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
 CC fibrosis, neurofibromatosis, tuberos sclerosis, Gaucher's disease,  
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary  
 CC haemostiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,  
 CC pulmonary alveolar proteinosis, Kargener syndrome, fibrocystic  
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension  
 CC and hyaline membrane disease. The present sequence is a peptide/protein  
 CC encoded by a single exon probe of the invention.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 29 AA;

Query Match 74.2%; Score 23; DB 23; Length 29;  
 Best Local Similarity 50.0%; Pred. No. 17;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXOEXD 10  
 | | | | |  
 DB 10 KALSTGOEID 19

RESULT 10  
 AAG63856  
 ID AAG63856 standard; Protein: 228 AA.  
 AC AAG63856;  
 XX  
 XX AAG63856;  
 DT 29-OCT-2001 (first entry)  
 DE Amino acid sequence of a modified Staphylococcal enterotoxin C1.  
 KW Enterotoxin C1; SEC-SER; Infectious disease; mastitis.  
 OS Synthetic.  
 OS Staphylococcus sp.  
 XX WO200160851-A1.  
 PN 23-AUG-2001.  
 PD 31-OCT-2000; 2000WO-KR01241.  
 PE 17-FEB-2000; 2000KR-0007612.  
 PR (GLDS ) LG CHEM LTD.  
 PA  
 XX Lee H, Park Y, Han K, Chang B, Lee Y;  
 PI WPI: 2001-522585/57.  
 DR N-PSDB; AAH74983.  
 DR  
 XX  
 PT Producing stable modified Staphylococcal toxin polypeptide for treating  
 PT infectious diseases, e.g. mastitis, in animals, involves substituting  
 PT cysteine at specified position of modified Staphylococcal toxin C1 with  
 PT serine -  
 PT  
 XX Claim 1; Page 59-60; 64pp; English.  
 PS  
 XX The present sequence represents a modified Staphylococcal enterotoxin C1,  
 CC designated SEC-SER. The modified toxin is characterised in that the  
 CC 95th amino acid (cysteine) is substituted with serine. This results in  
 CC a toxin that has improved stability. The modified enterotoxin is  
 CC useful for preventing or treating infectious diseases due to  
 CC microorganisms in animals such as cows, pigs, horses, sheep, hens,  
 CC dogs and cats, e.g. mastitis in cows.  
 CC  
 XX Sequence 228 AA;

Query Match 74.2%; Score 23; DB 22; Length 228;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXOEXD 10  
 | | | | |  
 DB 140 KKSVTAOEID 149

RESULT 11  
 AAR13208  
 ID AAR13208 standard; Protein: 238 AA.  
 AC AAR13208;  
 XX

DT 15-OCT-1991 (first entry)  
 XX Staphylococcal enterotoxin C3.  
 DE SEC3; cancer treatment; pyrogen; tumouricide.  
 KW Staphylococcus aureus.  
 OS WO9110680-A.  
 PN 25-JUL-1991.  
 PD 17-JAN-1991; 91WO-US00342.  
 PF 17-JAN-1990; 90US-0466577.  
 PR (TERM/) TERMAN D S;  
 PA Terman DS;  
 PI WPI: 1991-237984/32.  
 DR  
 XX  
 PT Treating cancer with enterotoxin from Staphylococcus aureus -  
 PT administered by IV injection, having same tumoricidal activity  
 PT as Staphylococcal protein A without potential toxic reactions  
 XX  
 PS Disclosure; Fig 1; 74pp; English.  
 XX SEC3 was isolated and purified from S.aureus. It can be used for  
 CC treating cancer, activating cytokine mediators and procoagulant  
 CC systems, augmenting natural killer cell cytotoxicity, etc. The  
 CC enterotoxin can be administered intravenously, optionally with  
 CC ibuprofen to attenuate toxic reaction to SEC3. Synthetic  
 CC polypeptides having structural homology to Staphylococcal exotoxins  
 CC are claimed, provided the homology includes statistically  
 CC significant sequence homology, alignment of cysteine residues and  
 CC similar hydrophathy profiles.  
 CC See AAR13203-R13211.  
 CC  
 XX Sequence 238 AA;

Query Match 74.2%; Score 23; DB 12; Length 238;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXOEXD 10  
 | | | | |  
 DB 150 KKSVTAOEID 159

RESULT 12  
 AAR45016  
 ID AAR45016 standard; Protein: 238 AA.  
 AC AAR45016;  
 XX  
 XX 08-JUN-1994 (first entry)  
 DT  
 XX Staphylococcal enterotoxin SEC3.  
 DE Staphylococcal enterotoxin; SE; cancer; tumouricidal agent;  
 KW autoimmune disease; toxicity; Protein A; perfusion system.  
 KW  
 XX Staphylococcus aureus.  
 OS WO9324136-A.  
 PN 09-DEC-1993.  
 PD 01-JUN-1993; 93WO-US05213.  
 PF 01-JUN-1992; 92US-0891718.  
 PR  
 XX

PA (STON/) STONE J L.  
 PA (TERM/) TERMAN D S.  
 XX Stone JL, Terman DS;  
 XX WPI: 1993-405418/50.  
 DR WPI: 1993-405418/50.  
 XX  
 XX Use of staphylococcal enterotoxin(s) and homologues - for  
 PT treating cancer in a patient or for the treatment of auto-immune  
 PT diseases  
 XX  
 XX Disclosure: Fig 1; 90pp; English.  
 XX  
 CC The sequences given in AAR45011-21 are Staphylococcal enterotoxins (SEs)  
 CC which may be used in the methods of the invention for treating cancer  
 CC in a patient. These SEs, and homologues of them, can be used as  
 CC tumoricidal agents for treating cancers and autoimmune disease.  
 CC They exhibit tumoricidal activity and toxicity identical to that  
 CC observed for the Protein A pertussis system. They may be administered  
 CC by i.v. injection.  
 CC  
 SQ Sequence 238 AA:  
 Query Match 74.2%; Score 23; DB 14; Length 238;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 OY 1 KXXXTXQEXD 10  
 DB 150 KKSXTAQLD 159  
 RESULT 13  
 AAB67343  
 ID AAB67343 standard; peptide; 238 AA.  
 XX AAB67343;  
 XX  
 DT 23-APR-2001 (first entry)  
 XX  
 DE Staphylococcus aureus enterotoxin C3 protein.  
 XX  
 KW Tumour; cancer; immune; enterotoxin.  
 OS Staphylococcus aureus.  
 XX  
 PN US6180097-B1.  
 PD 30-JAN-2001.  
 XX  
 PF 30-OCT-1998; 98US-0183437.  
 XX  
 PR 31-JAN-1994; 94US-0189424.  
 PR 19-JUN-1995; 95US-0491746.  
 PR 03-OCT-1989; 89US-0416530.  
 PR 17-JAN-1990; 90US-0466577.  
 PR 17-JAN-1991; 91WO-US00342.  
 PR 01-JUN-1992; 92US-0891718.  
 PR 02-MAR-1993; 93US-0025144.  
 XX  
 PA (TERM/) TERMAN D S.  
 PA  
 PI Terman DS;  
 XX  
 DR WPI: 2001-158657/16.  
 XX  
 PT Tumor cell capable of stimulating antitumor immune reactivity in vitro  
 PT or in vivo comprises exogenous nucleic acids encoding a superantigen  
 PT and a costimulatory molecule -  
 XX  
 XX Disclosure: Fig 2; 16pp; English.  
 XX  
 CC The present invention relates to a tumour cell capable of stimulating

CC antitumor immune reactivity in vitro or in vivo contains and  
 CC expresses an exogenous nucleic acid molecule encoding a superantigen  
 CC or its active fragment and an exogenous nucleic acid molecule  
 CC encoding a costimulatory molecule that activates T cells in  
 CC conjunction with an antigenic stimulus. The invention may be used  
 CC for cancer therapy by stimulating an anticancer immune response  
 CC in vivo or ex vivo.  
 CC  
 SQ Sequence 238 AA:  
 Query Match 74.2%; Score 23; DB 22; Length 238;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 OY 1 KXXXTXQEXD 10  
 DB 150 KKSXTAQLD 159  
 RESULT 14  
 ABB76239  
 ID ABB76239 standard; Protein; 238 AA.  
 XX ABB76239;  
 XX  
 DT 09-AUG-2002 (first entry)  
 XX  
 DE Staphylococcus aureus enterotoxin C3.  
 XX  
 KW Enterotoxin C3; SEC3; superantigen; antigen; tumour; cancer;  
 KW antitumour; therapy.  
 XX  
 OS Staphylococcus aureus.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 208 /note="given as 'O' in the specification"  
 PN US2002051765-A1.  
 PD 02-MAY-2002.  
 XX  
 PF 19-DEC-2000; 2000US-0741503.  
 XX  
 PR 31-JAN-1994; 94US-0189424.  
 PR 19-JUN-1995; 95US-0491746.  
 PR 03-OCT-1989; 89US-0416530.  
 PR 17-JAN-1990; 90US-0466577.  
 PR 17-JAN-1991; 91WO-US00342.  
 PR 01-JUN-1992; 92US-0891718.  
 PR 02-MAR-1993; 93US-0025144.  
 XX  
 PA (TERM/) TERMAN D S.  
 PA  
 PI Terman DS;  
 XX  
 DR WPI: 2002-415198/44.  
 XX  
 PT Reagent for treating cancer without the need for e.g. radiotherapy,  
 PT comprises a specific V beta subset of T cells sensitized to a growing  
 PT tumor and stimulated with superantigens -  
 XX  
 XX Disclosure: Fig 2; 17pp; English.  
 XX  
 CC The present sequence is the protein sequence of enterotoxin C3  
 CC (SEC3) of Staphylococcus aureus. Similarity is shown, in several  
 CC stretches of sequence, between staphylococcal enterotoxins,  
 CC streptococcal pyrogenic exotoxins and staphylococcal exfoliative  
 CC toxins (see ABB76234-44). In the present invention, synthetic  
 CC polypeptides useful in tumour therapy and in blocking or destroying  
 CC autoreactive T and B lymphocyte populations are characterised by  
 CC substantial structural homology to staphylococcal enterotoxin A and  
 CC enterotoxin B, and to streptococcal pyrogenic exotoxins, with

statistically significant sequence homology and similarity (Z value of Lipman and Pearson algorithm in Monte Carlo analysis exceeding 6) to include alignment of cysteine residues and similar hydrophathy profiles. These superantigens are used to treat solid tumors, including their metastases, without radiation, surgery or standard chemotherapeutic agents. A claimed method of human cancer treatment involves contacting haematopoietic cells from a patient with one or more superantigens ex vivo to generate stimulated cells, selecting a specific V beta subset of cells, and reintroducing these cells into the patient to induce an in vivo therapeutic, tumoricidal reaction.

Sequence 238 AA;

Query Match 74.2%; Score 23; DB 23; Length 238;  
Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
| | | |  
DB 150 KSVTAQELD 159

RESULT 15

AAR13207  
ID AAR13207 standard; Protein; 239 AA.

AC AAR13207;

DT 15-OCT-1991 (first entry)

DE Staphylococcal enterotoxin C1.

KW SEC1; cancer treatment; pyrogen; tumouricide.

OS Staphylococcus aureus.

FN W09110680-A.

PD 25-JUL-1991.

PF 17-JAN-1991; 91WO-US00342.

PR 17-JAN-1990; 90US-046577.

PA (TERM/) Terman D S.

PI Terman DS;

DR WPI: 1991-237984/32.

PT Treating cancer with enterotoxin from Staphylococcus aureus -  
administered by IV injection, having same tumoricidal activity  
as Staphylococcal protein A without potential toxic reactions

PS Disclosure: Fig 1; 74pp; English.

CC SEC1 was isolated and purified from S.aureus. It can be used for  
treating cancer, activating cytokine mediators and procoagulant  
systems, augmenting natural killer cell cytotoxicity, etc. The  
enterotoxin can be administered intravenously, optionally with  
albumin to attenuate toxic reaction to SEC1. Synthetic  
polypeptides having structural homology to staphylococcal exotoxins  
are claimed, provided the homology includes statistically  
significant sequence homology, alignment of Cysteine residues and  
similar hydrophathy profiles.  
See AAR1303-R13211.

SO Sequence 239 AA;

Query Match 74.2%; Score 23; DB 12; Length 239;  
Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
| | | |  
DB 151 KSVTAQELD 160

Search completed: June 5, 2003, 16:00:22  
Job time : 34.9394 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: June 5, 2003, 15:23:33 ; Search time 8.78788 Seconds  
(without alignments)  
33.481 Million cell updates/sec

Title: US-09-150-947f-13

Perfect score: 31

Sequence: 1 KXXXTXQEXD 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCRTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Dackfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	23	74.2	17	4	US-08-896-933-6
2	23	74.2	17	4	US-09-314-235-6
3	23	74.2	24	3	US-08-838-413A-20
4	23	74.2	238	4	US-08-896-933-28
5	23	74.2	238	4	US-09-314-235-28
6	23	74.2	239	4	US-08-896-933-27
7	23	74.2	239	4	US-09-314-235-27
8	23	74.2	266	4	US-09-144-776B-14
9	22	71.0	17	4	US-08-896-933-4
10	22	71.0	17	4	US-08-896-933-12
11	22	71.0	17	4	US-08-896-933-14
12	22	71.0	17	4	US-08-896-933-15
13	22	71.0	17	4	US-08-896-933-16
14	22	71.0	17	4	US-08-896-933-17
15	22	71.0	17	4	US-09-314-235-4
16	22	71.0	17	4	US-09-314-235-12
17	22	71.0	17	4	US-09-314-235-14
18	22	71.0	17	4	US-09-314-235-15
19	22	71.0	17	4	US-09-314-235-16
20	22	71.0	17	4	US-09-314-235-17
21	22	71.0	24	3	US-08-838-413A-19
22	22	71.0	24	3	US-08-838-413A-24
23	22	71.0	191	4	US-08-858-207A-381
24	22	71.0	221	4	US-08-896-933-29
25	22	71.0	221	4	US-09-314-235-29
26	22	71.0	239	4	US-08-896-933-26
27	22	71.0	239	4	US-09-314-235-26

28	22	71.0	239	4	US-09-144-776B-10	Sequence 10, Appl
29	22	71.0	251	4	US-09-144-776B-16	Sequence 16, Appl
30	22	71.0	255	1	US-08-446-918A-2	Sequence 2, Appl
31	22	71.0	255	2	US-08-580-806-2	Sequence 2, Appl
32	22	71.0	266	4	US-09-414-276-8	Sequence 8, Appl
33	22	71.0	266	4	US-09-144-776B-6	Sequence 6, Appl
34	22	71.0	266	4	US-09-190-824-2	Sequence 8, Appl
35	22	71.0	292	4	US-09-190-824-2	Sequence 8, Appl
36	22	71.0	922	2	US-08-464-402-2	Sequence 2, Appl
37	22	71.0	922	4	US-09-054-775C-2	Sequence 2, Appl
38	21	67.7	17	4	US-08-896-933-2	Sequence 2, Appl
39	21	67.7	17	4	US-08-896-933-8	Sequence 8, Appl
40	21	67.7	17	4	US-08-896-933-10	Sequence 10, Appl
41	21	67.7	17	4	US-09-314-235-2	Sequence 2, Appl
42	21	67.7	17	4	US-09-314-235-8	Sequence 8, Appl
43	21	67.7	17	4	US-09-314-235-10	Sequence 10, Appl
44	21	67.7	24	3	US-08-838-413A-4	Sequence 4, Appl
45	21	67.7	24	3	US-08-838-413A-18	Sequence 18, Appl

## ALIGNMENTS

```
RESULT 1
US-08-896-933-6
; Sequence 6, Application US/08896933
; Patent No. 6221351
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; FILE REFERENCE: 09629/005002
; CURRENT APPLICATION NUMBER: US/08/896,933
; EARLIER FILING DATE: 1997-07-18
; EARLIER APPLICATION NUMBER: 08/252,978
; NUMBER OF SEQ. ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ. ID NO 6
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-896-933-6
; Sequence 6, Application US/09314235
; Patent No. 6338845
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; FILE REFERENCE: 09629/005004
; CURRENT APPLICATION NUMBER: US/09/314,235
; EARLIER FILING DATE: 1997-07-18
; EARLIER APPLICATION NUMBER: 08/896,933
; EARLIER FILING DATE: 1997-07-18
; EARLIER APPLICATION NUMBER: 08/252,978
; EARLIER FILING DATE: 1994-06-02
; EARLIER APPLICATION NUMBER: 07/891,718
; EARLIER FILING DATE: 1992-06-01
; EARLIER APPLICATION NUMBER: US91/00342
; EARLIER FILING DATE: 1991-01-17
; EARLIER APPLICATION NUMBER: 07/466,577
```

Query Match 74.2%; Score 23; DB 4; Length 17;  
Best Local Similarity 50.0%; Pred. No. 0.86;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10  
DB 1 KKSVAQELD 10

EARLIER FILING DATE: 1990-01-17  
EARLIER APPLICATION NUMBER: 07/416,530  
EARLIER FILING DATE: 1989-10-03  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 6  
LENGTH: 17  
TYPE: PRT  
ORGANISM: Staphylococcus aureas  
US-09-314-235-6

Query Match 74.2% Score 23; DB 4; Length 17;  
Best Local Similarity 50.0%; Pred. No. 0.86;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXOEXD 10  
| | | | |  
DB 1 KSVTAQELD 10

RESULT 3  
US-08-838-413A-20  
Sequence 20, Application US/08838413A

GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: PEPTIDES USEFUL FOR  
REDUCING SYMPTOMS OF TOXIC SHOCK SYNDROME  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/838,413A  
FILING DATE: 07-APR-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: MORRY, MARY J.  
REGISTRATION NUMBER: 34,398  
REFERENCE/DOCKET NUMBER: 2016-4010  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)758-4800  
TELEFAX: (212)751-6849  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 24  
TYPE: AMINO ACID  
STRANDEDNESS: UNKNOWN  
TOPOLOGY: UNKNOWN  
MOLECULE TYPE: PEPTIDE  
US-08-838-413A-20

Query Match 74.2% Score 23; DB 3; Length 24;  
Best Local Similarity 50.0%; Pred. No. 1.3;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXOEXD 10  
| | | | |  
DB 1 KSVTAQELD 10

RESULT 4  
US-08-896-933-28  
Sequence 28, Application US/08896933

GENERAL INFORMATION:  
APPLICANT: Terman, David S.  
TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,  
SUPERANTIGENS, AND RELATED COMPOUNDS  
FILE REFERENCE: 09629/005002  
CURRENT APPLICATION NUMBER: US/08/896,933  
CURRENT FILING DATE: 1997-07-18  
EARLIER APPLICATION NUMBER: 08/252,978  
EARLIER FILING DATE: 1994-06-02  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 28  
LENGTH: 238  
TYPE: PRT  
ORGANISM: Staphylococcus aureas  
US-08-896-933-28

Query Match 74.2% Score 23; DB 4; Length 238;  
Best Local Similarity 50.0%; Pred. No. 17;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXOEXD 10  
| | | | |  
DB 150 KSVTAQELD 159

RESULT 5  
US-09-314-235-28  
Sequence 28, Application US/09314235

GENERAL INFORMATION:  
APPLICANT: Terman, David S.  
TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,  
SUPERANTIGENS, AND RELATED COMPOUNDS  
FILE REFERENCE: 09629/005004  
CURRENT APPLICATION NUMBER: US/09/314,235  
CURRENT FILING DATE: 1999-05-18  
EARLIER APPLICATION NUMBER: 08/896,933  
EARLIER FILING DATE: 1997-07-18  
EARLIER APPLICATION NUMBER: 08/252,978  
EARLIER FILING DATE: 1994-06-02  
EARLIER APPLICATION NUMBER: 07/891,718  
EARLIER FILING DATE: 1992-06-01  
EARLIER APPLICATION NUMBER: US91/00342  
EARLIER FILING DATE: 1991-01-17  
EARLIER APPLICATION NUMBER: 07/466,577  
EARLIER FILING DATE: 1990-01-17  
EARLIER APPLICATION NUMBER: 07/416,530  
EARLIER FILING DATE: 1989-10-03  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 28  
LENGTH: 238  
TYPE: PRT  
ORGANISM: Staphylococcus aureas  
US-09-314-235-28

Query Match 74.2% Score 23; DB 4; Length 238;  
Best Local Similarity 50.0%; Pred. No. 17;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXOEXD 10  
| | | | |  
DB 150 KSVTAQELD 159

RESULT 6



```
US-08-896-933-27
; Sequence 27, Application US/08896933
; Patent No. 6221351
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; FILE REFERENCE: 09629/005002
; SUPERANTIGENS, AND RELATED COMPOUNDS
; CURRENT FILING DATE: 1997-07-18
; EARLIER FILING DATE: 1994-06-02
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27
; LENGTH: 239
; TYPE: PRF
; ORGANISM: Staphylococcus aureas
US-08-896-933-27

Query Match          74.2%; Score 23; DB 4; Length 239;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10
DB 151 KKSVAQELD 160

RESULT 7
US-09-314-235-27
; Sequence 27, Application US/09314235
; Patent No. 6338845
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; FILE REFERENCE: 09629/005004
; SUPERANTIGENS, AND RELATED COMPOUNDS
; CURRENT FILING DATE: 1999-05-18
; EARLIER FILING DATE: 1997-07-18
; EARLIER APPLICATION NUMBER: 08/896,933
; EARLIER FILING DATE: 1994-06-02
; EARLIER APPLICATION NUMBER: 07/891,718
; EARLIER FILING DATE: 1992-06-01
; EARLIER APPLICATION NUMBER: US91/00342
; EARLIER FILING DATE: 1991-01-17
; EARLIER APPLICATION NUMBER: 07/466,577
; EARLIER FILING DATE: 1990-01-17
; EARLIER APPLICATION NUMBER: 07/416,530
; EARLIER FILING DATE: 1989-10-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27
; LENGTH: 239
; TYPE: PRF
; ORGANISM: Staphylococcus aureas
US-09-314-235-27

Query Match          74.2%; Score 23; DB 4; Length 239;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10
DB 151 KKSVAQELD 160

RESULT 8
US-09-144-776B-14
; Sequence 14, Application US/09144776B
; Patent No. 6399332
```

```
GENERAL INFORMATION:
APPLICANT: Robert G. Ulrich,
Mark A. Olson
Sina Bavari
TITLE OF INVENTION: Bacterial Superantigen
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Charles H. Harris
STREET: US Army MRC -504 Scott Street
MCMR-JA (Charles H. Harris-Patent
Acty)
CITY: FORT DETRICK
STATE: MARYLAND
COUNTRY: USA
ZIP: 21702-5012
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.5
SOFTWARE: Microsoft Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/144,776B
FILING DATE: 01-Sep-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/882,431
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Charles H. Harris
REGISTRATION NUMBER: 34,616
REFERENCE/DOCKET NUMBER: <Unknown>
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 619-2065
TELEFAX: (301) 619-7714
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 266
TYPE: Amino Acid
STRANDEDNESS: Unknown
MOLECULE TYPE: Peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-144-776B-14

Query Match          74.2%; Score 23; DB 4; Length 266;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10
DB 178 KKSVAQELD 187

RESULT 9
US-08-896-933-4
; Sequence 4, Application US/08896933
; Patent No. 6221351
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; FILE REFERENCE: 09629/005002
; SUPERANTIGENS, AND RELATED COMPOUNDS
; CURRENT FILING DATE: 1997-07-18
; EARLIER FILING DATE: 1994-06-02
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 17
; TYPE: PRF
; ORGANISM: Staphylococcus aureas
```

US-08-896-933-4

Query Match 71.0%; Score 22; DB 4; Length 17;  
Best Local Similarity 50.0%; Pred. No. 1.7;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10  
DB 1 KKNVTAQELD 10

RESULT 10

US-08-896-933-12  
Sequence 12, Application US/08896933  
Patent No. 6221351  
GENERAL INFORMATION:  
APPLICANT: Terman, David S.  
TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,  
FILE REFERENCE: 09629/005002  
CURRENT APPLICATION NUMBER: US/08/896,933  
CURRENT FILING DATE: 1997-07-18  
EARLIER APPLICATION NUMBER: 08/252,978  
EARLIER FILING DATE: 1994-06-02  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 12  
LENGTH: 17  
TYPE: PRT  
ORGANISM: Streptococcus pyogenes  
US-08-896-933-12

Query Match 71.0%; Score 22; DB 4; Length 17;  
Best Local Similarity 50.0%; Pred. No. 1.7;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10  
DB 1 KKNVTAQELD 10

RESULT 11

US-08-896-933-14  
Sequence 14, Application US/08896933  
Patent No. 6221351  
GENERAL INFORMATION:  
APPLICANT: Terman, David S.  
TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,  
FILE REFERENCE: 09629/005002  
CURRENT APPLICATION NUMBER: US/08/896,933  
CURRENT FILING DATE: 1997-07-18  
EARLIER APPLICATION NUMBER: 08/252,978  
EARLIER FILING DATE: 1994-06-02  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 14  
LENGTH: 17  
TYPE: PRT  
ORGANISM: Other  
FEATURE:  
OTHER INFORMATION: Consensus sequences derived from staphylococcus  
; OTHER INFORMATION: aureus and streptococcus pyogenes polypeptides  
US-08-896-933-14

Query Match 71.0%; Score 22; DB 4; Length 17;  
Best Local Similarity 50.0%; Pred. No. 1.7;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10  
DB 1 KKNVTAQELD 10

RESULT 12

US-08-896-933-15  
Sequence 15, Application US/08896933  
Patent No. 6221351  
GENERAL INFORMATION:  
APPLICANT: Terman, David S.  
TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,  
FILE REFERENCE: 09629/005002  
CURRENT APPLICATION NUMBER: US/08/896,933  
CURRENT FILING DATE: 1997-07-18  
EARLIER APPLICATION NUMBER: 08/252,978  
EARLIER FILING DATE: 1994-06-02  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 15  
LENGTH: 17  
TYPE: PRT  
ORGANISM: Other  
FEATURE:  
OTHER INFORMATION: Consensus sequences derived from staphylococcus  
; OTHER INFORMATION: aureus and streptococcus pyogenes polypeptides  
US-08-896-933-15

Query Match 71.0%; Score 22; DB 4; Length 17;  
Best Local Similarity 50.0%; Pred. No. 1.7;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10  
DB 1 KKNVTAQELD 10

RESULT 13

US-08-896-933-16  
Sequence 16, Application US/08896933  
Patent No. 6221351  
GENERAL INFORMATION:  
APPLICANT: Terman, David S.  
TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,  
FILE REFERENCE: 09629/005002  
CURRENT APPLICATION NUMBER: US/08/896,933  
CURRENT FILING DATE: 1997-07-18  
EARLIER APPLICATION NUMBER: 08/252,978  
EARLIER FILING DATE: 1994-06-02  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 16  
LENGTH: 17  
TYPE: PRT  
ORGANISM: Other  
FEATURE:  
OTHER INFORMATION: Consensus sequences derived from staphylococcus  
; OTHER INFORMATION: aureus and streptococcus pyogenes polypeptides  
US-08-896-933-16

Query Match 71.0%; Score 22; DB 4; Length 17;  
Best Local Similarity 50.0%; Pred. No. 1.7;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10  
DB 1 KKNVTAQELD 10

RESULT 14

US-08-896-933-17  
Sequence 17, Application US/08896933  
Patent No. 6221351  
GENERAL INFORMATION:  
APPLICANT: Terman, David S.

;; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,  
;; TITLE OF INVENTION: SUPERANTIGENS, AND RELATED COMPOUNDS  
;; FILE REFERENCE: 09629/005002  
;; CURRENT APPLICATION NUMBER: US/08/896,933  
;; CURRENT FILING DATE: 1997-07-18  
;; EARLIER APPLICATION NUMBER: 08/252,978  
;; EARLIER FILING DATE: 1994-06-02  
;; NUMBER OF SEQ ID NOS: 34  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 17  
;; LENGTH: 17  
;; TYPE: PRT  
;; ORGANISM: Other  
;; FEATURE:  
;; OTHER INFORMATION: Consensus sequences derived from staphylococcus  
;; OTHER INFORMATION: aureus and streptococcus pyogenes polypeptides  
US-08-896-933-17

Query Match 71.0%; Score 22; DB 4; Length 17;  
Best Local Similarity 50.0%; Pred. No. 1.7;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10  
| | | | |  
Db 1 KKNVTAGELD 10

RESULT 15  
US-09-314-235-4  
;; Sequence 4, Application US/09314235  
;; Patent No. 6338845  
;; GENERAL INFORMATION:  
;; APPLICANT: Teerman, David S.  
;; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,  
;; TITLE OF INVENTION: SUPERANTIGENS, AND RELATED COMPOUNDS  
;; FILE REFERENCE: 09629/005004  
;; CURRENT APPLICATION NUMBER: US/09/314,235  
;; CURRENT FILING DATE: 1999-05-18  
;; EARLIER APPLICATION NUMBER: 08/896,933  
;; EARLIER FILING DATE: 1997-07-18  
;; EARLIER APPLICATION NUMBER: 08/252,978  
;; EARLIER FILING DATE: 1994-06-02  
;; EARLIER APPLICATION NUMBER: 07/891,718  
;; EARLIER FILING DATE: 1992-06-01  
;; EARLIER APPLICATION NUMBER: US91/00342  
;; EARLIER FILING DATE: 1991-01-17  
;; EARLIER APPLICATION NUMBER: 07/466,577  
;; EARLIER FILING DATE: 1990-01-17  
;; EARLIER APPLICATION NUMBER: 07/416,530  
;; EARLIER FILING DATE: 1989-10-03  
;; NUMBER OF SEQ ID NOS: 34  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 4  
;; LENGTH: 17  
;; TYPE: PRT  
;; ORGANISM: Staphylococcus aureus  
US-09-314-235-4

Query Match 71.0%; Score 22; DB 4; Length 17;  
Best Local Similarity 50.0%; Pred. No. 1.7;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10  
| | | | |  
Db 1 KKNVTAGELD 10

Search completed: June 5, 2003, 15:32:42  
Job time : 9.78788 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:31:01 ; Search time 12.1212 Seconds  
(without alignments)  
85.173 Million cell updates/sec

Title: US-09-150-947F-13  
Perfect score: 31  
Sequence: 1 KXXXTXQEXD 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	74.2	29	10	US-09-864-761-43487
2	23	74.2	266	1	US-08-882-431-14
3	23	74.2	266	9	US-09-870-759-12
4	23	74.2	266	9	US-10-002-784A-14
5	22	71.0	10	10	US-09-150-947B-3
6	22	71.0	10	10	US-09-150-947B-4
7	22	71.0	12	9	US-10-113-809-1
8	22	71.0	12	10	US-10-113-809-4
9	22	71.0	12	10	US-09-150-947B-2
10	22	71.0	13	10	US-09-150-947B-5
11	22	71.0	13	10	US-09-150-947B-6
12	22	71.0	13	10	US-09-150-947B-11
13	22	71.0	14	10	US-09-150-947B-9
14	22	71.0	14	10	US-09-150-947B-10
15	22	71.0	16	10	US-09-150-947B-7
16	22	71.0	26	10	US-09-150-947B-1
17	22	71.0	35	10	US-09-150-947B-8
18	22	71.0	102	9	US-10-047-260-26
19	22	71.0	220	9	US-10-002-784A-26

20	22	71.0	239	1	US-08-882-431-10	Sequence 10, Appl
21	22	71.0	239	9	US-10-002-784A-10	Sequence 10, Appl
22	22	71.0	239	10	US-09-150-947B-12	Sequence 12, Appl
23	22	71.0	251	1	US-08-882-431-16	Sequence 16, Appl
24	22	71.0	251	8	US-08-973-391A-13	Sequence 13, Appl
25	22	71.0	251	9	US-10-002-784A-16	Sequence 16, Appl
26	22	71.0	265	1	US-08-882-431-6	Sequence 6, Appl
27	22	71.0	265	1	US-08-882-431-8	Sequence 6, Appl
28	22	71.0	266	9	US-09-870-759-10	Sequence 10, Appl
29	22	71.0	266	9	US-10-002-784A-6	Sequence 6, Appl
30	22	71.0	266	9	US-10-002-784A-8	Sequence 8, Appl
31	22	71.0	266	9	US-10-151-336-8	Sequence 8, Appl
32	22	71.0	295	10	US-09-815-242-13351	Sequence 1351, A
33	22	71.0	295	10	US-09-815-242-13691	Sequence 13691, A
34	22	71.0	344	9	US-10-201-310-3	Sequence 3, Appl
35	22	71.0	350	9	US-09-972-473-38	Sequence 38, Appl
36	22	71.0	468	9	US-10-002-784A-27	Sequence 27, Appl
37	22	71.0	659	9	US-10-102-627-74	Sequence 74, Appl
38	22	71.0	922	10	US-09-879-228-2	Sequence 2, Appl
39	21	67.7	12	9	US-10-113-809-2	Sequence 2, Appl
40	21	67.7	72	9	US-10-144-929-189	Sequence 189, App
41	21	67.7	177	9	US-10-092-849-5	Sequence 5, Appl
42	21	67.7	177	10	US-09-730-617-95	Sequence 95, Appl
43	21	67.7	190	9	US-09-738-626-4637	Sequence 4637, Ap
44	21	67.7	203	9	US-09-900-766-5	Sequence 5, Appl
45	21	67.7	217	9	US-09-900-766-6	Sequence 6, Appl

## ALIGNMENTS

RESULT 1  
US-09-864-761-43487  
Sequence 43487, Application US/09864761  
Patent No. US2002048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecomla-X-1  
CURRENT APPLICATION NUMBER: US/09/864, 761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263,6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 43487  
LENGTH: 29  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC006299.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.8  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.7  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.67  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.7  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.66  
OTHER INFORMATION: EST\_HUMAN HIT: AV703706.1, EVALU6 6.00e+00  
US-09-864-761-43487

Query Match  
Best Local Similarity 74.2%; Score 23; DB 10; Length 29;  
Matches 5; Conservative 0; Pred. No. 6.8; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXOEXD 10  
DB 10 KALSTGEID 19

RESULT 2  
US-08-882-431-14  
Sequence 14, Application US/08882431  
Publication No. US20030009015A1  
GENERAL INFORMATION:  
APPLICANT: Robert G. Ulrich,  
APPLICANT: Mark A. Olson  
APPLICANT: Sina Bavari  
TITLE OF INVENTION: Bacterial Superantigen  
TITLE OF INVENTION: Vaccines  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John Moran  
STREET: US Army MRC -504 Scott Street MCMR-JA (John Moran-Patent Atty)  
CITY: FORT DERRICK  
STATE: MARYLAND  
COUNTRY: USA  
ZIP: 21702-5012  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh 7.5  
SOFTWARE: Microsoft Word 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/882,431  
FILING DATE: June 25, 1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Moran, John  
REGISTRATION NUMBER: 26,313  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 619-2065  
TELEFAX: (301) 619-7714  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 266

TYPE: Amino Acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
US-08-882-431-14

Query Match  
Best Local Similarity 74.2%; Score 23; DB 1; Length 266;  
Matches 5; Conservative 0; Pred. No. 71; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXOEXD 10  
DB 178 KKSVTAGEID 187

RESULT 3  
US-09-870-759-12  
Sequence 12, Application US/09870759  
Patent No. US20020177551A1  
GENERAL INFORMATION:  
APPLICANT: TERMAN, David S  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE  
FILE REFERENCE: 870759  
CURRENT APPLICATION NUMBER: US/09/870,759  
CURRENT FILING DATE: 2002-01-14  
PRIOR APPLICATION NUMBER: US 60/208,128  
PRIOR FILING DATE: 2000-05-30  
NUMBER OF SEQ ID NOS: 166  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 12  
LENGTH: 266  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-870-759-12

Query Match  
Best Local Similarity 74.2%; Score 23; DB 9; Length 266;  
Matches 5; Conservative 0; Pred. No. 71; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXOEXD 10  
DB 178 KKSVTAGEID 187

RESULT 4  
US-10-002-784A-14  
Sequence 14, Application US/10002784A  
Publication No. US20030036644A1  
GENERAL INFORMATION:  
/33  
APPLICANT: Ulrich, Robert G.  
TITLE OF INVENTION: Bacterial Superantigen Vaccines  
FILE REFERENCE: 003/233/SAP  
CURRENT APPLICATION NUMBER: US/10/002,784A  
CURRENT FILING DATE: 2001-11-26  
PRIOR APPLICATION NUMBER: 08/882,431; 09/144,776  
PRIOR FILING DATE: 97-06-25; 98-09-01  
NUMBER OF SEQ ID NOS: 40  
SOFTWARE: Apple Macintosh Microsoft Word 6.0  
SEQ ID NO 14  
LENGTH: 266  
TYPE: PRT  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: staphylococcal enterotoxin C-1 mutant  
US-10-002-784A-14

Query Match  
Best Local Similarity 74.2%; Score 23; DB 9; Length 266;  
Matches 5; Conservative 0; Pred. No. 71; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXOEXD 10  
DB 178 KKSVTAGEID 187

Fri-Jun 6 09:14:26 2003

Db 178 KKSVTAQELD 187

## RESULT 5

US-09-150-947B-3

; Sequence 3, Application US/09150947B  
; Patent No. US20020028211A1

; GENERAL INFORMATION:

; APPLICANT: Kaempfer, Raymond

; APPLICANT: Atad, Gila

; TITLE OF INVENTION: BROAD SPECTRUM PYROGENIC EXOTOXINS

; FILE REFERENCE: A31967-PCT-USA-A

; CURRENT APPLICATION NUMBER: US/09/150,947B

; CURRENT FILING DATE: 1998-09-10

; PRIOR APPLICATION NUMBER: PCT/IL97/00438

; PRIOR FILING DATE: 1997-12-30

; PRIOR APPLICATION NUMBER: ISRAEL 119938

; PRIOR FILING DATE: 1996-12-30

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 3

; LENGTH: 10

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-150-947B-3

Query Match 71.0%; Score 22; DB 10; Length 10;

Best Local Similarity 50.0%; Pred. No. 4.2;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTQXEND 10

Db 1 KKKVTAQELD 10

## RESULT 6

US-09-150-947B-4

; Sequence 4, Application US/09150947B  
; Patent No. US20020028211A1

; GENERAL INFORMATION:

; APPLICANT: Kaempfer, Raymond

; APPLICANT: Atad, Gila

; TITLE OF INVENTION: BROAD SPECTRUM PYROGENIC EXOTOXINS

; FILE REFERENCE: A31967-PCT-USA-A

; CURRENT APPLICATION NUMBER: US/09/150,947B

; CURRENT FILING DATE: 1998-09-10

; PRIOR APPLICATION NUMBER: PCT/IL97/00438

; PRIOR FILING DATE: 1997-12-30

; PRIOR APPLICATION NUMBER: ISRAEL 119938

; PRIOR FILING DATE: 1996-12-30

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 4

; LENGTH: 10

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-150-947B-4

Query Match 71.0%; Score 22; DB 10; Length 10;

Best Local Similarity 50.0%; Pred. No. 4.2;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTQXEND 10

Db 1 KKKVTAQELD 10

## RESULT 7

US-10-113-809-1

; Sequence 1, Application US/10113809  
; Patent No. US2002017753A1

; GENERAL INFORMATION:

; APPLICANT: Kaempfer, Raymond

; APPLICANT: Atad, Gila

; TITLE OF INVENTION: BROAD SPECTRUM PYROGENIC EXOTOXINS

; FILE REFERENCE: A31967-PCT-USA-A

; CURRENT APPLICATION NUMBER: US/09/150,947B

; CURRENT FILING DATE: 1998-09-10

; PRIOR APPLICATION NUMBER: PCT/IL97/00438

; PRIOR FILING DATE: 1997-12-30

; PRIOR APPLICATION NUMBER: ISRAEL 119938

; GENERAL INFORMATION:

; APPLICANT: Maschke, Hans, E.

; TITLE OF INVENTION: EXOTOXIN-LIGAND

; FILE REFERENCE: MBP-011XX

; CURRENT APPLICATION NUMBER: US/10/113,809

; CURRENT FILING DATE: 2002-03-29

; PRIOR APPLICATION NUMBER: 101 16 042.9-41

; PRIOR FILING DATE: 2001-03-30

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 12

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Bacterial Toxin Binding Ligand

US-10-113-809-1

Query Match 71.0%; Score 22; DB 9; Length 12;

Best Local Similarity 50.0%; Pred. No. 5.1;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTQXEND 10

Db 3 KKKVTAQELD 12

## RESULT 8

US-10-113-809-4

; Sequence 4, Application US/10113809  
; Patent No. US2002017753A1

; GENERAL INFORMATION:

; APPLICANT: Maschke, Hans, E.

; TITLE OF INVENTION: EXOTOXIN-LIGAND

; FILE REFERENCE: MBP-011XX

; CURRENT APPLICATION NUMBER: US/10/113,809

; CURRENT FILING DATE: 2002-03-29

; PRIOR APPLICATION NUMBER: 101 16 042.9-41

; PRIOR FILING DATE: 2001-03-30

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 12

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Bacterial Toxin Binding Ligand

US-10-113-809-4

Query Match 71.0%; Score 22; DB 9; Length 12;

Best Local Similarity 50.0%; Pred. No. 5.1;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTQXEND 10

Db 3 KKKVTAQELD 12

## RESULT 9

US-09-150-947B-2

; Sequence 2, Application US/09150947B  
; Patent No. US20020028211A1

; GENERAL INFORMATION:

; APPLICANT: Kaempfer, Raymond

; APPLICANT: Atad, Gila

; TITLE OF INVENTION: BROAD SPECTRUM PYROGENIC EXOTOXINS

; FILE REFERENCE: A31967-PCT-USA-A

; CURRENT APPLICATION NUMBER: US/09/150,947B

; CURRENT FILING DATE: 1998-09-10

; PRIOR APPLICATION NUMBER: PCT/IL97/00438

; PRIOR FILING DATE: 1997-12-30

; PRIOR APPLICATION NUMBER: ISRAEL 119938

;; PRIOR FILING DATE: 1996-12-30  
;; NUMBER OF SEQ ID NOS: 12  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 2  
;; LENGTH: 12  
;; TYPE: PRT  
;; ORGANISM: Staphylococcus aureus  
US-09-150-947B-2

Query Match 71.0%; Score 22; DB 10; Length 12;  
Best Local Similarity 50.0%; Pred.No. 5.1;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10  
| | | | |  
DB 3 KKKATVQELD 12

RESULT 10  
US-09-150-947B-5  
; Sequence 5, Application US/09150947B  
; Patent No. US20020028211A1  
; GENERAL INFORMATION:  
; APPLICANT: Kaempfer, Raymond  
; APPLICANT: Arad, Gila  
; TITLE OF INVENTION: BROAD SPECTRUM PYROGENIC EXOTOXINS  
; TITLE OF INVENTION: ANTAGONISTS AND VACCINES  
; FILE REFERENCE: A31967-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/150,947B  
; CURRENT FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: PCT/IL97/00438  
; PRIOR FILING DATE: 1997-12-30  
; PRIOR APPLICATION NUMBER: 1997-12-30  
; PRIOR FILING DATE: 1996-12-30  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 5  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
; FEATURE:  
; NAME/KEY: LIPID  
; LOCATION: (1)...(1)  
; OTHER INFORMATION: N-lauryl cysteine residue  
US-09-150-947B-5

Query Match 71.0%; Score 22; DB 10; Length 13;  
Best Local Similarity 50.0%; Pred.No. 5.5;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10  
| | | | |  
DB 4 KKKYTAQELD 13

RESULT 11  
US-09-150-947B-6  
; Sequence 6, Application US/09150947B  
; Patent No. US20020028211A1  
; GENERAL INFORMATION:  
; APPLICANT: Kaempfer, Raymond  
; APPLICANT: Arad, Gila  
; TITLE OF INVENTION: BROAD SPECTRUM PYROGENIC EXOTOXINS  
; TITLE OF INVENTION: ANTAGONISTS AND VACCINES  
; FILE REFERENCE: A31967-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/150,947B  
; CURRENT FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: PCT/IL97/00438  
; PRIOR FILING DATE: 1997-12-30  
; PRIOR APPLICATION NUMBER: ISRAEL 119938  
; PRIOR FILING DATE: 1996-12-30  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 3.0

;; SEQ ID NO 6  
;; LENGTH: 13  
;; TYPE: PRT  
;; ORGANISM: Staphylococcus aureus  
;; FEATURE:  
;; NAME/KEY: LIPID  
;; LOCATION: (1)...(1)  
;; OTHER INFORMATION: N-lauryl cysteine residue  
US-09-150-947B-6

Query Match 71.0%; Score 22; DB 10; Length 13;  
Best Local Similarity 50.0%; Pred.No. 5.5;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10  
| | | | |  
DB 4 KKKATVQELD 13

RESULT 12  
US-09-150-947B-11  
; Sequence 11, Application US/09150947B  
; Patent No. US20020028211A1  
; GENERAL INFORMATION:  
; APPLICANT: Kaempfer, Raymond  
; APPLICANT: Arad, Gila  
; TITLE OF INVENTION: BROAD SPECTRUM PYROGENIC EXOTOXINS  
; TITLE OF INVENTION: ANTAGONISTS AND VACCINES  
; FILE REFERENCE: A31967-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/150,947B  
; CURRENT FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: PCT/IL97/00438  
; PRIOR FILING DATE: 1997-12-30  
; PRIOR APPLICATION NUMBER: ISRAEL 119938  
; PRIOR FILING DATE: 1996-12-30  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 11  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
; FEATURE:  
; NAME/KEY: ACETYLATION  
; LOCATION: (1)...(1)  
; OTHER INFORMATION: N-acetyl group  
; NAME/KEY: MOD\_RES  
; LOCATION: (13)...(13)  
; OTHER INFORMATION: D-alanine  
US-09-150-947B-11

Query Match 71.0%; Score 22; DB 10; Length 13;  
Best Local Similarity 50.0%; Pred.No. 5.5;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10  
| | | | |  
DB 3 KKKATVQELD 12

RESULT 13  
US-09-150-947B-9  
; Sequence 9, Application US/09150947B  
; Patent No. US20020028211A1  
; GENERAL INFORMATION:  
; APPLICANT: Kaempfer, Raymond  
; APPLICANT: Arad, Gila  
; TITLE OF INVENTION: BROAD SPECTRUM PYROGENIC EXOTOXINS  
; TITLE OF INVENTION: ANTAGONISTS AND VACCINES  
; FILE REFERENCE: A31967-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/150,947B  
; CURRENT FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: PCT/IL97/00438  
; PRIOR FILING DATE: 1997-12-30



```
; PRIOR APPLICATION NUMBER: ISRAEL 119938
; PRIOR FILING DATE: 1996-12-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: DISULFID
; LOCATION: (1)...(1)
; NAME/KEY: DISULFID
; LOCATION: (14)...(14)
US-09-150-947B-9
```

```
Query Match      71.0%; Score 22; DB 10; Length 14;
Best Local Similarity 50.0%; Pred. No. 6;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 KXXXTXQEXD 10
      | | | |
DB      4 KKKATVQELD 13
```

```
RESULT 14
US-09-150-947B-10
; Sequence 10, Application US/09150947B
; Patent No. US2002028211A1
; GENERAL INFORMATION:
; APPLICANT: Kaempfer, Raymond
; APPLICANT: Arad, Gila
; TITLE OF INVENTION: BROAD SPECTRUM PYROGENIC EXOTOXINS
; FILE REFERENCE: A31967-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/150,947B
; CURRENT FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: PCT/IL97/00438
; PRIOR FILING DATE: 1997-12-30
; PRIOR APPLICATION NUMBER: ISRAEL 119938
; PRIOR FILING DATE: 1996-12-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (1)...(1)
; OTHER INFORMATION: D-alanine
; NAME/KEY: MOD_RES
; LOCATION: (14)...(14)
; OTHER INFORMATION: D-alanine
US-09-150-947B-10
```

```
Query Match      71.0%; Score 22; DB 10; Length 14;
Best Local Similarity 50.0%; Pred. No. 6;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 KXXXTXQEXD 10
      | | | |
DB      4 KKKATVQELD 13
```

```
RESULT 15
US-09-150-947B-7
; Sequence 7, Application US/09150947B
; Patent No. US2002028211A1
; GENERAL INFORMATION:
; APPLICANT: Kaempfer, Raymond
; APPLICANT: Arad, Gila
; TITLE OF INVENTION: BROAD SPECTRUM PYROGENIC EXOTOXINS
; TITLE OF INVENTION: ANTAGONISTS AND VACCINES
```

```
; FILE REFERENCE: A31967-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/150,947B
; CURRENT FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: PCT/IL97/00438
; PRIOR FILING DATE: 1997-12-30
; PRIOR APPLICATION NUMBER: ISRAEL 119938
; PRIOR FILING DATE: 1996-12-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-150-947B-7
```

```
Query Match      71.0%; Score 22; DB 10; Length 16;
Best Local Similarity 50.0%; Pred. No. 6.9;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 KXXXTXQEXD 10
      | | | |
DB      3 KKKATVQELD 12
```

```
Search completed: June 5, 2003, 15:49:02
Job time : 13.1212 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:24:01 ; Search time 9.84848 seconds  
(without alignments)  
97.613 Million cell updates/sec

Title: US-09-150-947F-13

Perfect score: 31

Sequence: 1 KXXXTXQEXD 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	77.4	549	2 E86337	hypothetical prote
2	24	77.4	602	1 RRNZP3	polymerase-associ
3	24	77.4	603	1 RRNZP5	polymerase-associ
4	24	77.4	604	1 RRNZP4	polymerase-associ
5	23	74.2	266	1 ENSAC1	enterotoxin C-1 pr
6	23	74.2	266	2 A60114	enterotoxin C-2 pr
7	23	74.2	266	2 S11885	enterotoxin C3 - S
8	23	74.2	565	2 A87694	ribosomal protein
9	23	74.2	566	2 AB3061	30S ribosomal prot
10	23	74.2	568	1 R3ZR1	ribosomal protein
11	23	74.2	572	1 E98225	30S ribosomal prot
12	23	74.2	586	1 ORXLE	estrogen receptor
13	23	74.2	3212	2 T24692	hypothetical prote
14	22	71.0	102	2 S74843	hypothetical prote
15	22	71.0	136	2 A89969	enterotoxin YEMT2
16	22	71.0	184	2 F86681	prophage pil prote
17	22	71.0	236	2 S18789	exotoxin A precurs
18	22	71.0	236	2 S18786	exotoxin type A pr
19	22	71.0	236	2 S18783	exotoxin type A pr
20	22	71.0	239	2 D89969	enterotoxin SEM [1
21	22	71.0	240	2 G89991	extracellular ente
22	22	71.0	242	2 C89969	extracellular ente
23	22	71.0	247	2 T22335	hypothetical prote
24	22	71.0	251	1 S29659	exotoxin type A pr
25	22	71.0	258	1 QOECF	Replication initia
26	22	71.0	260	2 H89968	enterotoxin Sen [1
27	22	71.0	260	2 E89969	enterotoxin Sen [1
28	22	71.0	266	1 ENSAB6	enterotoxin B prec
29	22	71.0	295	2 D95107	hypothetical prote

30	22	71.0	295	2 F97975	hypothetical prote
31	22	71.0	332	1 DEBYG2	glyceraldhyde-3-P
32	22	71.0	333	2 JC5787	ribonuclease H (EC
33	22	71.0	356	2 T38408	hypothetical prote
34	22	71.0	520	2 G71647	hypothetical prote
35	22	71.0	523	2 F97865	hypothetical prote
36	22	71.0	672	2 F71424	hypothetical prote
37	22	71.0	835	2 T05259	probable disease r
38	22	71.0	886	2 T39229	hypothetical prote
39	22	71.0	922	2 I37292	DNA ligase (ATP) (
40	22	71.0	946	2 A84133	hypothetical prote
41	22	71.0	1146	2 H96796	hypothetical prote
42	22	71.0	1200	2 S77524	chromosome segrega
43	22	71.0	1314	2 A85176	hypothetical prote
44	22	71.0	1327	2 D70759	probable ocsb prot
45	22	71.0	1738	2 T14867	interactin - slime

## ALIGNMENTS

## RESULT 1

E86337 hypothetical protein F14010.3 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001

C:Accession: E86337

R:Rheologis, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alon

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Hultzer, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

C.A.; Li, J.H.; Li, Y.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maitl, R.; Marzla

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: E86337

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1549 <STO>

A:Cross-references: GB:AEO05172; NID:99558589; PIDN:AAF88152.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

Query Match	77.4%	Score 24;	DB 2;	Length 549;
Best Local Similarity	50.0%	Pred. No. 62;		
Matches	5;	Conservative	0;	Mismatches
			5;	Indels
				Gaps
				0;
Qy	1	KXXXTXQEXD 10		
Db	27	KISTGQEND 36		

## RESULT 2

RRNZP3 polymerase-associated nucleocapsid phosphoprotein - parainfluenza virus type 3 (strat

C:Species: parainfluenza virus type 3

C>Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 24-Jul-1997

C:Accession: A27010

R:Spriggs, M.K.; Collins, P.L.

J. Gen. Virol. 67, 2705-2719, 1986

A:Title: Sequence analysis of the P and C protein genes of human parainfluenza virus

A:Reference number: A92793; MUID:87085488; PMID:3025346

A:Accession: A27010

A:Molecule type: mRNA

A:Residues: 1-602 <SPR>

C:Comment: This protein may be a component of the active polymerase.

C:Genetics:

A:Gene: P

C:Superfamily: parainfluenza virus polymerase-associated nucleocapsid phosphoprotein

C:Keywords: nucleocapsid; phosphoprotein

Query Match 77.4%; Score 24; DB 1; Length 602;  
Best Local Similarity 50.0%; Pred. NO. 68;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
DB 226 KSSSTHQEDD 235

#### RESULT 3

RNNZ5  
polymerase-associated nucleocapsid phosphoprotein (version 2) - parainfluenza virus type

C:Species: parainfluenza virus type 3  
A:Note: host: Homo sapiens (man)

C>Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 24-Jul-1997

C:Accession: A26896

R:Galinski, M.S.; Mink, M.A.; Lambert, D.M.; Wechsler, S.L.; Pons, M.W.  
Virology 155, 46-60, 1986

A:Title: Molecular cloning and sequence analysis of the human parainfluenza 3 virus mRNA

A:Reference number: A94348; MUID:87044104; PMID:3022477

A:Accession: A26896

A:Molecule type: mRNA

A:Residues: 1-603 <GAL>

C:Comment: The RNA sequence was obtained from GenBank, release 52.0.

C:Genetics: This protein may be a component of the active polymerase.

A:Gene: P

C:Superfamily: parainfluenza virus polymerase-associated nucleocapsid phosphoprotein

C:Keywords: nucleocapsid; phosphoprotein

Query Match 77.4%; Score 24; DB 1; Length 603;  
Best Local Similarity 50.0%; Pred. NO. 68;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
DB 226 KSSSTHQEDD 235

#### RESULT 4

RNNZ4  
polymerase-associated nucleocapsid phosphoprotein (version 1) - parainfluenza virus type

C:Species: parainfluenza virus type 3  
A:Note: host: Homo sapiens (man)

C>Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 24-Jul-1997

C:Accession: A94355; A24189

R:Luk, D.; Sanchez, A.; Banerjee, A.K.  
Virology 156, 193-194, 1987

A:Reference number: A94355

A:Contents: extratum

A:Accession: A94355

A:Molecule type: mRNA

A:Residues: 1-604 <LUI>

A:Cross-references: GB:M14890

R:Luk, D.; Sanchez, A.; Banerjee, A.K.  
Virology 153, 318-325, 1986

A:Title: Messenger RNA encoding the phosphoprotein (P) gene of human parainfluenza virus

A:Reference number: A94343; MUID:86291173; PMID:3016995

A:Accession: A24189

A:Molecule type: mRNA

A:Residues: 1-603; AKDQIKTTPNK' <LUI>

A:Note: this sequence has been corrected in reference A94355

C:Comment: This protein may be a component of the active polymerase.

C:Genetics: This protein may be a component of the active polymerase.

A:Gene: P

C:Superfamily: parainfluenza virus polymerase-associated nucleocapsid phosphoprotein

C:Keywords: nucleocapsid; phosphoprotein

Query Match 77.4%; Score 24; DB 1; Length 604;  
Best Local Similarity 50.0%; Pred. NO. 69;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
DB 226 KSSSTHQEDD 235

#### RESULT 5

ENSAC1  
enterotoxin C-1 precursor - Staphylococcus aureus

C:Species: Staphylococcus aureus

C>Date: 15-Nov-1984 #sequence\_revision 05-Jan-1996 #text\_change 18-Jun-1999

C:Accession: S06356; A01816

R:Bohach, G.A.; Schlievert, P.M.

Mol. Gen. Genet. 209, 15-20, 1987

A:Title: Nucleotide sequence of the staphylococcal enterotoxin C1 gene and relatedness

A:Reference number: S06356; MUID:88038352; PMID:2823067

A:Accession: S06356

A:Molecule type: DNA

A:Residues: 1-266 <BOH>

A:Cross-references: EMBL:X05815; NID:g46566; PIDN:CAA29260.1; PID:g46567

R:Schmidt, J.J.; Spero, L.

J. Biol. Chem. 258, 6300-6306, 1983

A:Title: The complete amino acid sequence of Staphylococcal enterotoxin C-1.

A:Reference number: A01816; MUID:83213327; PMID:6189824

A:Accession: A01816

A:Molecule type: protein

A:Residues: 28-75, 'IL', '78-176, 'N', '178-266 <SCH>

C:Genetics:

A:Gene: entC1

C:Superfamily: enterotoxin B

C:Keywords: enterotoxin

F:1-27/Domain: signal sequence #status predicted <SIG>

F:28-266/Product: enterotoxin C-1 #status experimental <MAT>

F:120-137/Disulfide bonds: #status experimental

Query Match 74.2%; Score 23; DB 1; Length 266;  
Best Local Similarity 50.0%; Pred. NO. 55;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
DB 178 KKSVAQELD 187

RESULT 6  
A60114  
enterotoxin C-2 precursor - Staphylococcus aureus

N:Alternate names: enterotoxin C-3 precursor

C:Species: Staphylococcus aureus

C>Date: 10-Nov-1992 #sequence\_revision 10-Nov-1992 #text\_change 16-Jul-1999

C:Accession: A60114; B60114; A33866

R:Bohach, G.A.; Schlievert, P.M.

Infect. Immun. 57, 2249-2252, 1989

A:Title: Conservation of the biologically active portions of staphylococcal enterotox

A:Reference number: A60114; MUID:89277549; PMID:2543637

A:Accession: A60114

A:Molecule type: DNA

A:Residues: 1-266 <BOH>

A:Accession: B60114

A:Molecule type: protein

A:Residues: 28-66 <BOH2>

R:Couch, J.L.; Bentley, M.J.

J. Bacteriol. 171, 4507-4510, 1989

A:Title: Nucleotide sequence of the type C-3 staphylococcal enterotoxin gene suggests

A:Reference number: A33866; MUID:89327174; PMID:2473979

A:Accession: A33866

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-266 <COU>

A:Cross-references: GB:M28364; NID:g153003; PIDN:AAA26624.1; PID:g153004

C:Genetics:

A:Gene: entC2

C:Superfamily: enterotoxin B  
F:1-27/Domain: signal sequence #status predicted <SIG>  
F:28-266/Product: enterotoxin C-2 #status predicted <MAT>

Query Match 74.2% Score 23; DB 2; Length 266;  
Best Local Similarity 50.0%; Pred. No. 55;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
DB 178 KKSXTAQELD 187

## RESULT 7

S11885 enterotoxin C3 - Staphylococcus aureus

C:Species: Staphylococcus aureus  
C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 16-Jul-1999

C:Accession: S11885  
R:Hoove, C.J.; Hackett, S.P.; Bohach, G.A.  
Mol. Gen. Genet. 220, 329-333, 1990

A:Title: Nucleotide sequence of the staphylococcal enterotoxin C3 gene: sequence comparison  
A:Reference number: S11885; MID:90220508; PMID:2325627

A:Accession: S11885  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-266 <HOV>  
A:Cross-references: GB:AE00689; MID:88217521; PMID:3368316  
C:Superfamily: enterotoxin B

Query Match 74.2% Score 23; DB 2; Length 266;  
Best Local Similarity 50.0%; Pred. No. 55;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
DB 178 KKSXTAQELD 187

## RESULT 8

A87694 ribosomal protein S1 [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus  
C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001  
C:Accession: A87694

R:Nierman, W.C.; Debby, R.T.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.B.; Laub, M.T.; Doherty, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon, P.; Ermlaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; MID:21173698; PMID:11259647

A:Accession: A87694  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-565 <STO>  
A:Cross-references: GB:AE005673; MID:913425331; PIDN:AAK25549.1; GSPDB:GN00148  
C:Gene: CC3587  
C:Superfamily: Escherichia coli ribosomal protein S1

Query Match 74.2% Score 23; DB 2; Length 565;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
DB 326 KIVSTSQEVD 335

## RESULT 9

AB3061 30S ribosomal protein S1 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C:Species: Agrobacterium tumefaciens

C>Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 01-Feb-2002  
C:Accession: AB3061  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Moe, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayvin, T.; Levy, R.; Li, M.; McCl, Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamster, B.W.

A:Title: The genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AB3061

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-566 <KUR>

A:Cross-references: GB:AE00689; PIDN:AAI44904.1; PID:917742555; GSPDB:GN00187

A:Experimental source: strain C58 (Dupont)

C:Gene: rpsA

A:Gene: rpsA

A:Map position: linear chromosome

C:Superfamily: Escherichia coli ribosomal protein S1

Query Match 74.2% Score 23; DB 2; Length 566;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
DB 326 KILSTQEVD 335

## RESULT 10

R32R1 ribosomal protein S1 - Rhizobium meliloti

C:Species: Rhizobium meliloti  
C>Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 22-Jun-1999  
C:Accession: S01055  
R:Schlier, J.; Thamm, S.; Lutz, R.; Hussain, A.; Faist, G.; Dobrinski, B.

Nucleic Acids Res. 16, 3075-3089, 1988

A:Title: Cloning and characterization of a gene from Rhizobium meliloti 2011 coding

A:Reference number: S01055; MID:88217521; PMID:3368316

A:Accession: S01055

A:Molecule type: DNA

A:Residues: 1-568 <SCH>

A:Cross-references: EMBL:X07528; MID:946338; PIDN:CAA30404.1; PID:946339

A>Note: the sequence from Fig. 5 is inconsistent with that from Fig. 4 in having 565-

C:Gene: rpsA

A:Gene: rpsA

C:Superfamily: Escherichia coli ribosomal protein S1

C:Keywords: duplication; protein biosynthesis; ribosome; RNA binding

F:196-453/Domain: RNA binding #status predicted <RNA>

F:196-279,280-366,367-453/Region: duplication

Query Match 74.2% Score 23; DB 1; Length 568;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
DB 326 KILSTQEVD 335

## RESULT 11

E98225 30S ribosomal protein S1 [imported] - Agrobacterium tumefaciens (strain C58, Cereon)

C:Species: Agrobacterium tumefaciens  
C>Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 11-Jan-2002

C:Accession: E98225

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ouello, B.; Goldm

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium

A:Reference number: A97359; PMID:11743194

A:Accession: E98225

A:Map position: linear chromosome

A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-572 <KUR>  
 A:Cross-references: GB:AE007870; PIDN:AAK89327.1; PID:g15159167; GSPDB:GN00170  
 C:Genetics:  
 A:Gene: AGR.L.1496  
 A:Map position: linear chromosome  
 C:Superfamily: Escherichia coli ribosomal protein S1

Query Match 74.2% Score 23; DB 2; Length 572;  
 Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
 DB 332 KILSTQEXD 341

# RESULT 12

ORXLE  
 estrogen receptor - African clawed frog  
 C:Species: Xenopus laevis (African clawed frog)  
 C:Date: 28-Feb-1992 #sequence\_revision 14-Jul-1994 #text\_change 15-Jun-1996  
 C:Accession: A40907  
 R:Weller, J.J.; Lew, D.; Shapiro, D.J.  
 Mol. Endocrinol. 1, 355-362, 1987  
 A:Title: The Xenopus laevis estrogen receptor: sequence homology with human and avian re  
 A:Reference number: A40907; MUID:90331927; PMID:3274894  
 A:Accession: A40907  
 A:Molecule type: mRNA  
 A:Residues: 1-586 <MEI>

A:Cross-references: GB:L20735  
 C:Comment: The steroid hormones and their receptors are involved in the regulation of eu  
 C:Comment: In the absence of ligand, steroid hormone receptors are thought to be weakly  
 C:Comment: complex appears to recognize discrete DNA sequences upstream of transcriptional start sit  
 C:Superfamily: estrogen receptor; erba transforming protein homology  
 C:Keywords: DNA binding; nucleus; phosphoprotein; steroid binding; steroid hormone recep  
 F:1-174/Domain: amino-terminal <NH2>  
 F:175-266/Domain: DNA binding #status predicted <DNA>  
 F:178-448/Domain: erba transforming protein homology <ERBA>  
 F:178-201/Region: zinc finger CCCC motif  
 F:214-237/Region: zinc finger CCCC motif  
 F:251-266/Region: nuclear location signal  
 F:292-544/Domain: steroid binding #status predicted <STB>  
 F:180,183,197,200/Binding site: zinc (Cys) #status predicted  
 F:216,222,232,235/Binding site: zinc (Cys) #status predicted  
 F:231/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 74.2% Score 23; DB 1; Length 586;  
 Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
 DB 545 KDKTTQEXD 554

# RESULT 13

T24692  
 hypothetical protein T08G11.1 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T24692  
 R:Dobson, R.  
 submitted to the EMBL Data Library, September 1996  
 A:Reference number: Z19925  
 A:Accession: T24692

A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-3212 <WIL>  
 A:Cross-references: EMBL:Z80220; PIDN:CAB02304.1; GSPDB:GN00019; CESP:T08G11.1  
 A:Experimental source: clone T08G11  
 C:Genetics:

A:Gene: CESP:T08G11.1  
 A:Map position: 1  
 A:Introns: 34/1; 149/3; 182/3; 527/2; 766/1; 1070/3; 1769/3; 1834/3; 2070/2; 2210/3;

Query Match 74.2% Score 23; DB 2; Length 3212;  
 Best Local Similarity 50.0%; Pred. No. 8e+02;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
 DB 2264 KSVTQEXD 2273

# RESULT 14

S74843  
 hypothetical protein s110846 - Synechocystis sp. (strain PCC 6803)  
 C:Species: Synechocystis sp.  
 A:Variety: PCC 6803  
 C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
 C:Accession: S74843  
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima,  
 O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas  
 DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys  
 A:Reference number: S74322; MUID:97061201; PMID:8905231  
 A:Accession: S74843  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-102 <KAN>  
 A:Cross-references: EMBL:D90909; GB:AB001339; NID:g1653844; PIDN:BA17804.1; PID:g101  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 71.0% Score 22; DB 2; Length 102;  
 Best Local Similarity 50.0%; Pred. No. 38;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
 DB 72 KKKTKQEXD 81

# RESULT 15

A89969  
 enterotoxin YEMT2 [imported] - Staphylococcus aureus (strain N315)  
 C:Species: Staphylococcus aureus  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
 C:Accession: A89969  
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O  
 ma, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K  
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
 Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
 A:Reference number: A89758; MUID:21311952; PMID:11418146  
 A:Accession: A89969  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-136 <KUR>  
 A:Cross-references: GB:BA000018; PID:g13701619; PIDN:BA042912.1; GSPDB:GN00149  
 A:Experimental source: strain N315  
 C:Genetics:

Query Match 71.0% Score 22; DB 2; Length 136;  
 Best Local Similarity 50.0%; Pred. No. 51;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
 DB 50 KKNITQEXD 59

Search completed: June 5, 2003, 15:33:57

Fri Jun 6 09:14:29 2003

Job time : 13.8485 secs

us-09-150-947f-13.rpr

Page 5

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:34:07 ; Search time 6.51515 Seconds  
(without alignments)  
63.661 Million cell updates/sec

Title: US-09-150-947F-13

Perfect score: 31

Sequence: 1 KXXXTXQEXD 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	77.4	604	1 RRRP_P13H4	P06162 human para
2	74.2	266	1 ETC1_STAU	P01553 staphylococ
3	74.2	266	1 ETC2_STAU	P34071 staphylococ
4	74.2	266	1 ETC3_STAU	P23313 staphylococ
5	74.2	568	1 RSL_RHIME	P14129 rhizobium m
6	74.2	586	1 ESRL_XENLA	P81559 xenopus lae
7	71.0	251	1 REEL_ECOLI	P03856 escherichia
8	71.0	251	1 SPEA_STRPY	P08095 streptococ
9	71.0	266	1 ETCB_STAU	P01552 staphylococ
10	71.0	331	1 G3P3_YEAST	P00359 saccharomyc
11	71.0	350	1 DKK3_CHICK	P00839 gallus gall
12	71.0	356	1 YDGC_SCHPO	Q09039 schizosacch
13	71.0	520	1 Y855_RICPR	Q92435 rickettsia
14	71.0	712	1 ARS2_CABEL	Q96615 caenorhabdl
15	71.0	886	1 YEB6_SCHPO	Q14302 schizosacch
16	71.0	922	1 DNL3_HUMAN	P49916 homo sapien
17	71.0	1015	1 DNL3_MOUSE	P97386 mus musculu
18	71.0	1327	1 YK06_MYCTU	Q10850 mycobacteri
19	67.7	176	1 YRS9_LISMO	Q84583 listeria mo
20	67.7	177	1 ILIX_HORSE	Q18899 equus cabal
21	67.7	228	1 GLI_ARATH	P27900 arabidopsis
22	67.7	234	1 SPEG_STRPY	Q9X5C7 streptococ
23	67.7	235	1 SPEC_STRPY	P13380 streptococ
24	67.7	236	1 SPEH_STRPY	Q9X5C8 streptococ
25	67.7	257	1 ETCX_STAU	P13163 staphylococ
26	67.7	257	1 ETCX_STAU	P12993 staphylococ
27	67.7	258	1 ETCX_STAU	P20723 staphylococ
28	67.7	258	1 ETCX_STAU	O85382 staphylococ
29	67.7	276	1 PYGI_SYNEL	P50039 synchococc
30	67.7	283	1 YJ81_YEAST	P47148 saccharomyc
31	67.7	351	1 PEKE_PICAN	P78723 pichia angu
32	67.7	375	1 NP4A_HUMAN	O99733 homo sapien
33	67.7	385	1 TGT_ZYMO	P28720 zymomonas m

34	21	67.7	461	1 DISK_MOUSE	Q61466 mus musculu
35	21	67.7	480	1 SYP_MYCLE	Q92517 mycobacteri
36	21	67.7	554	1 NBL4_MOUSE	P52963 mus musculu
37	21	67.7	684	1 FLID_HELPY	Q92191 helicobacte
38	21	67.7	684	1 FLID_HELPY	P96786 helicobacte
39	21	67.7	727	1 GPM_HUMAN	P43304 homo sapien
40	21	67.7	727	1 GPM_MOUSE	Q64521 mus musculu
41	21	67.7	899	1 ANDR_MOUSE	P19091 mus musculu
42	21	67.7	904	1 N1A1_TOBAC	P11605 nicotiana t
43	21	67.7	904	1 N1A2_TOBAC	P08509 nicotiana t
44	21	67.7	1184	1 XMS2_DROME	Q94399 drosophila
45	21	67.7	1305	1 GAK_RAT	P97874 rattus norv

#### ALIGNMENTS

```

RESULT 1
ID RRRP_P13H4 STANDARD; PRT; 604 AA.
AC P06162;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE RNA polymerase alpha subunit (EC 2.7.7.48) (Nucleocapsid
DE phosphoprotein).
GN P.
OS Human parainfluenza 3 virus (strain NIH 47885).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Respirovirus.
OX NCBI_TaxID=11217;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86291173; PubMed=3016995;
RA Luk D., Sanchez A., Banerjee A.K.;
RT "Messenger RNA encoding the phosphoprotein (P) gene of human
RT parainfluenza virus 3 is bicistronic.";
RL Virology 153:318-325(1986).
RN [2]
RP REVISIONS.
RA Luk D., Sanchez A., Banerjee A.K.;
RL Virology 156:193-194(1987).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=87044104; PubMed=3022477;
RA Gallinski M.S., Mink M.A., Lambert D.M., Wechsler S.L., Pons W.M.;
RT "Molecular cloning and sequence analysis of the human parainfluenza 3
RT virus mRNA encoding the P and C proteins.";
RL Virology 155:46-60(1986).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=87085488; PubMed=3025346;
RA Spriggs M.K., Collins P.L.;
RT "Sequence analysis of the P and C protein genes of human
RT parainfluenza virus type 3: patterns of amino acid sequence homology
RT among paramyxovirus proteins.";
RL J. Gen. Virol. 67:2705-2719(1986).
CC -1- FUNCTION: THIS PROTEIN IS PROBABLY A COMPONENT OF THE ACTIVE
CC POLYMERASE. IT MAY FUNCTION IN TEMPLATE BINDING.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC {RNA}(n).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M14890; AAA6866.1; ALT-SEQ.
CC DR X04721; CAA28429.1; -
CC EMBL: M14932; AAA6818.1; -

```

DR EMBL: D10029; BA00921.1; ALT\_SEQ.  
 DR PIR: A27010; RRN2P3.  
 DR PIR: A24189; RRN2P4.  
 DR PIR: A26896; RRN2P5.  
 DR InterPro: IPR002693; Paramyxo\_P.  
 DR Pfam: PF01806; Paramyxo\_P.1.  
 KM Transferase; RNA-directed RNA polymerase: Nucleocapsid;  
 KM Phosphorylation.  
 FT CONFLICT 198 198 P -> N (IN REF. 4).  
 FT CONFLICT 243 243 MISSING (IN REF. 3 AND 4).  
 FT CONFLICT 244 244 MISSING (IN REF. 4).  
 FT CONFLICT 316 316 Q -> R (IN REF. 3).  
 FT CONFLICT 584 584 C -> S (IN REF. 4).  
 SQ SEQUENCE 604 AA; 67661 MW; D2E06EF47027D0BD CRC64;

Query Match 77.4%; Score 24; DB 1; Length 604;  
 Best Local Similarity 50.0%; Pred. No. 25;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
 DB 226 KSSSTHQEDD 235

## RESULT 2

ETC1\_STRAU STANDARD; PRT; 266 AA.  
 AC P01553;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Enterotoxin type C-1 precursor (SEC1).  
 GN EMTC1.  
 OS Staphylococcus aureus.  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OX NCBI\_TaxID=1280;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88038352; PubMed=2823067;  
 RA Bohach G.A., Schlievert P.M.;  
 RT "Nucleotide sequence of the staphylococcal enterotoxin C1 gene and  
 RL relatedness to other pyrogenic toxins.";  
 RL Mol. Gen. Genet. 209:15-20(1987).  
 RN [2]  
 RP SEQUENCE OF 28-266.  
 RX MEDLINE=83213327; PubMed=6189824;  
 RA Schmidt J.J., Spero L.;  
 RT "The complete amino acid sequence of staphylococcal enterotoxin C1.";  
 RL J. Biol. Chem. 258:6300-6306(1983).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION  
 CC STAPHYLOCOCCAL FOOD POISONING SYNDROME.  
 CC -1- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN  
 CC FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: X05815; CAA29260.1; -  
 DR PIR: A01816; ENSACL.  
 DR PIR: S06356; S06356.  
 DR HSSP: P34071; 1SE2.  
 DR InterPro: IPR001961; Staph/Strep\_toxin.  
 DR Pfam: PF01123; Staph\_stp\_toxin; 1.  
 DR Pfam: PF02876; Staph\_stp\_tox\_C; 1.  
 DR PRINTS: PR00279; BACTRTOXIN.  
 DR PROSITE: PS00277; STAPH\_STREP\_TOXIN\_1; 1.  
 DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.

KW Enterotoxin; Toxin; Signal; Superantigen.  
 FT SIGNAL 1 27  
 FT CHAIN 28 266 ENTEROTOXIN TYPE C-1.  
 FT DISULFID 120 137  
 FT CONFLICT 177 177 D -> N (IN REF. 2).  
 SQ SEQUENCE 266 AA; 30546 MW; 3A7AB59A898653B CRC64;

Query Match 74.2%; Score 23; DB 1; Length 266;  
 Best Local Similarity 50.0%; Pred. No. 20;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
 DB 178 KSVTAQELD 187

## RESULT 3

ETC2\_STRAU STANDARD; PRT; 266 AA.  
 AC P34071;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Enterotoxin type C-2 precursor (SEC2).  
 GN EMTC2.  
 OS Staphylococcus aureus.  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OX NCBI\_TaxID=1280;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-66.  
 RX MEDLINE=89277549; PubMed=2543637;  
 RA Bohach G.A., Schlievert P.M.;  
 RT "Conservation of the biologically active portions of staphylococcal  
 RL enterotoxins C1 and C2.";  
 RL Infect. Immun. 57:2249-2252(1989).  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
 RX MEDLINE=96027099; PubMed=7582894;  
 RA Papageorgiou A.C., Acharya K.R., Shapiro R., Passalacqua E.F.,  
 RA Brehm R.D., Tranter H.S.;  
 RT "Crystal structure of the superantigen enterotoxin C2 from  
 RL staphylococcus aureus reveals a zinc-binding site.";  
 RL Structure 3:769-779(1995).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).  
 RX MEDLINE=96022987; PubMed=7552730;  
 RA Swaminathan S., Furey W.F. Jr., Pletcher J., Sax M.;  
 RT "Residues defining V beta specificity in staphylococcal  
 RL enterotoxins.";  
 RL Nat. Struct. Biol. 2:680-686(1995).  
 RN [4]  
 RP COMPARISON OF STRUCTURE OF SEA AND SEC2.  
 RX MEDLINE=97334373; PubMed=9191070.  
 RA Schad E.M., Papageorgiou A.C., Svensson L.A., Acharya K.R.;  
 RT "A structural and functional comparison of staphylococcal  
 RL enterotoxins A and C2 reveals remarkable similarity and  
 RL dissimilarity.";  
 RL J. Mol. Biol. 269:270-280(1997).  
 CC -1- COFACTOR: Binds 1 zinc ion per subunit. The zinc ion is necessary  
 CC for the toxin interaction with MHC class II.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION  
 CC STAPHYLOCOCCAL FOOD POISONING SYNDROME.  
 CC -1- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN  
 CC FAMILY.  
 CC  
 DR PIR: A60114; A60114.  
 DR PDB: 1STE; 23-DEC-96.  
 DR PDB: 1SE2; 08-MAR-96.  
 DR InterPro: IPR001961; Staph/Strep\_toxin.  
 DR Pfam: PF01123; Staph\_stp\_toxin; 1.  
 DR Pfam: PF02876; Staph\_stp\_tox\_C; 1.  
 DR PRINTS: PR00279; BACTRTOXIN.  
 DR PROSITE: PS00277; STAPH\_STREP\_TOXIN\_1; 1.

DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.  
 KW Enterotoxin; Toxin; Signal; Superantigen; Metal-binding; Zinc;  
 3D-structure.  
 FT SIGNAL 1 27  
 FT CHAIN 28 266 ENTEROTOXIN TYPE C-2.  
 FT DISULFID 120 137  
 FT METAL 36 36 ZINC.  
 FT METAL 110 110 ZINC.  
 FT METAL 145 145 ZINC.  
 FT METAL 149 149 ZINC.  
 SQ SEQUENCE 266 AA; 30604 MW; 8407FB18536FAC08 CRC64;

Query Match 74.2%; Score 23; DB 1; Length 266;  
 Best Local Similarity 50.0%; Pred. No. 20;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEND 10  
 DB 178 KKSVAQELD 187

RESULT 4  
 ETC3\_STPAM STANDARD; PRT; 266 AA.  
 ID ETC3\_STPAM  
 AC P23313;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Enterotoxin type C-3 precursor (SEC3).  
 GN ETC3 OR SA2009 OR SA1817.  
 OS Staphylococcus aureus (strain Mu50 / ATCC 700699),  
 OS Staphylococcus aureus (strain N315), and  
 OS Staphylococcus aureus.  
 OC Bacteria: Firmicutes: Bacillales: Staphylococcus.  
 OX NCBI\_TaxID=158878, 158879, 1280;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Mu50 / ATCC 700699, and N315;  
 RX MEDLINE=21311952; PubMed=11418146;  
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,  
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-Q., Ito T.,  
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,  
 RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,  
 RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,  
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,  
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;  
 RA "Whole genome sequencing of methicillin-resistant Staphylococcus  
 aureus";  
 RT Lancet 357:1225-1240(2001).  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90220508; PubMed=2325627;  
 RA Hovde C.J., Hackett S.P., Bohach G.A.;  
 RT "Nucleotide sequence of the staphylococcal enterotoxin C3 gene:  
 sequence comparison of all three type C staphylococcal  
 enterotoxins";  
 RT Mol. Gen. Genet. 220:329-333(1990).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS) OF COMPLEX WITH TCR.  
 RX MEDLINE=97064178; PubMed=8906797;  
 RA Fields B.A., Malchiodi E.L., Li H., Yeern X., Stauffacher C.V.,  
 RA Schlievert P.M., Karjalainen K., Marluza R.A.;  
 RT "Crystal structure of a T-cell receptor beta-chain complexed with a  
 superantigen";  
 RT Nature 384:188-192(1996).  
 RL [4]  
 RP SUBCELLULAR LOCATION: Secreted.  
 CC -1- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION  
 CC STAPHYLOCOCCAL FOOD POISONING SYNDROME.  
 CC -1- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN  
 CC FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AP003364; BAB58171.1; -  
 CC EMBL: AP003135; BAB43097.1; -  
 CC EMBL: X51661; CA35972.1; -  
 CC PIR: S11885; S11885.  
 CC PDB: 1JCK; 12-NOV-97.  
 DR InterPro: IPR001961; Staph/Strep\_toxin.  
 DR Pfam: PF01123; Staph\_Strep\_toxin; 1.  
 DR Pfam: PF02876; Staph\_Strep\_toxin; 1.  
 DR PRINTS: PR00279; BACTRTOXIN.  
 DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_1; 1.  
 DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.  
 KW Enterotoxin; Toxin; Signal; Superantigen; 3D-structure;  
 Complete proteome.  
 FT SIGNAL 1 27  
 FT CHAIN 28 266 ENTEROTOXIN TYPE C-3.  
 FT DISULFID 120 137  
 SQ SEQUENCE 266 AA; 30671 MW; 5ED8A32D1FFCA59 CRC64;

Query Match 74.2%; Score 23; DB 1; Length 266;  
 Best Local Similarity 50.0%; Pred. No. 20;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEND 10  
 DB 178 KKSVAQELD 187

RESULT 5  
 RSL\_RHIME STANDARD; PRT; 568 AA.  
 ID RSL\_RHIME  
 AC P14129;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 30S ribosomal protein S1.  
 GN RPSA OR R00255 OR SMC00335.  
 OS Rhizobium meliloti (Sinorhizobium meliloti).  
 OC Bacteria: Proteobacteria: alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae: Sinorhizobium.  
 OX NCBI\_TaxID=382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-RCR2011 / S047;  
 RX MEDLINE=88217521; PubMed=3368316;  
 RA Schindler J., Thamm S., Lutz R., Hussain A., Falst G.,  
 RA Dobrinski B.;  
 RT "Cloning and characterization of a gene from Rhizobium meliloti 2011  
 coding for ribosomal protein S1";  
 RT Nucleic Acids Res. 16:3075-3089(1988).  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-1021;  
 RX MEDLINE=21396507; PubMed=11481430;  
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,  
 RA Bolstad P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
 RA Gohl T., Goffeau A., Kahn D., Kiss E., Leleure V., Masuy D.,  
 RA Pohl T., Portetalle D., Puehler A., Purnelle B., Ransperger U.,  
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Gallibert F.;  
 RT "Analysis of the chromosome sequence of the legume symbiont  
 Sinorhizobium meliloti strain 1021";  
 RT Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
 CC -1- FUNCTION: BINDS MRNA; THUS FACILITATING RECOGNITION OF THE  
 CC INITIATION POINT. IT IS NEEDED TO TRANSLATE MRNA WITH A SHORT  
 CC SHINE-DALGARNO (SD) PURINE-RICH SEQUENCE (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE S1P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC -1- SIMILARITY: CONTAINS 6 SI MOTIF DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL; X07528; CAA30404.1; -  
 DR EMBL; AL591783; CAC41692.1; -  
 DR PIR; S01055; R3ZRL.  
 DR HSSP; P05055; ISRO.  
 DR InterPro; IPR000110; Ribosomal\_S1.  
 DR InterPro; IPR003029; S1.  
 DR Pfam; PF00575; S1; 6.  
 DR PRINTS; PR00681; RIBOSOMAL\_S1.  
 DR SMART; SM00316; S1; 6.  
 DR TIGRfam; TIGR00717; rpsA; 1.  
 DR PROSITE; PS50126; S1; 6.  
 KW Ribosomal protein; Repeat; RNA-binding; Complete proteome.  
 FT DOMAIN 27 93 S1 MOTIF 1.  
 FT DOMAIN 111 177 S1 MOTIF 2.  
 FT DOMAIN 198 266 S1 MOTIF 3.  
 FT DOMAIN 283 353 S1 MOTIF 4.  
 FT DOMAIN 370 440 S1 MOTIF 5.  
 FT DOMAIN 459 530 S1 MOTIF 6.  
 FT CONFLICT 88 A -> R (IN REF. 1).  
 FT CONFLICT 153 162 LMHNPPPEI -> ADAPPALRN (IN REF. 1).  
 FT CONFLICT 235 235 N -> K (IN REF. 1).  
 FT CONFLICT 241 241 L -> Q (IN REF. 1).  
 FT CONFLICT 551 551 A -> R (IN REF. 1).  
 SQ SEQUENCE 568 AA; 62640 MW; AB0858204273A7B8 CRC64;

Query Match 74.2%; Score 23; DB 1; Length 568;  
 Best Local Similarity 50.0%; Pred. No. 45;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10  
 DB 326 KILSTQEVLD 335

RESULT 6  
 ID ESRL\_XENLA STANDARD; PRT; 586 AA.  
 AC P81559;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Estrogen receptor (ER) (Estradiol receptor) (ER-alpha).  
 GN ESRL OR NR3A1 OR ESR.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
 CC Xenopodinae; Xenopus.  
 OC NCBI\_Taxid=8335;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90331927; PubMed=3274894;  
 RA Weller I.J., Lew D., Shapiro D.J.;  
 RT "The Xenopus laevis estrogen receptor: sequence homology with human  
 RT and avian receptors and identification of multiple estrogen receptor  
 RT messenger ribonucleic acids."  
 RL Mol. Endocrinol. 1:355-362(1987).  
 CC -1- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED  
 CC IN THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT  
 CC CELLULAR PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.  
 CC -1- SUBUNIT: BINDS DNA AS A HOMODIMER. CAN FORM A HETERODIMER WITH ER-  
 CC BETA (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,  
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.

CC NR3 SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL; L20735; -; NOT\_ANNOTATED\_CDS.  
 DR HSSP; P03372; HRCQ.  
 DR TRANSFAC; T00263; -  
 DR InterPro; IPR000536; Hormone\_rec\_1lg.  
 DR InterPro; IPR001292; Oestrgn\_receptor.  
 DR InterPro; IPR001723; Sthhrm\_receptor.  
 DR InterPro; IPR001628; Znf\_Casteroid.  
 DR Pfam; PF00104; hormone\_rec; 1.  
 DR Pfam; PF00105; Zf-C4; 1.  
 DR Pfam; PF02159; Oest\_rec; 1.  
 DR PRINTS; PR00398; STRDHOMONER.  
 DR PRODOM; PD000035; Znf\_Casteroid; 1.  
 DR SMART; SM00430; HOL1; 1.  
 DR SMART; SM00399; Znf\_C4; 1.  
 DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
 KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;  
 KW Zinc-finger; Steroid-binding.  
 FT DOMAIN 1 179 MODULATING.  
 FT DNA\_BIND 180 245 NUCLEAR RECEPTOR-TYPE.  
 FT ZN\_FING 180 200 C4-TYPE.  
 FT ZN\_FING 216 240 C4-TYPE.  
 FT DOMAIN 246 302 HINGE.  
 FT DOMAIN 303 586 STEROID-BINDING.  
 SQ SEQUENCE 586 AA; 66080 MW; 0EDCC7EBDD6F08BF CRC64;

Query Match 74.2%; Score 23; DB 1; Length 586;  
 Best Local Similarity 50.0%; Pred. No. 46;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10  
 DB 545 KDKTTQEDD 554

RESULT 7  
 ID REEL\_ECOLI STANDARD; PRT; 251 AA.  
 AC P03856; P08969;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Replication initiation protein (RIP protein) (Protein E) (F4 protein).  
 GN REPE OR REP OR E.  
 OS Escherichia coli.  
 OC Plasmid F.  
 CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 CC Escherichia.  
 OC NCBI\_Taxid=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=82051318; PubMed=7028572;  
 RA Murotsu T., Matsubara K., Sugisaki H., Takanami M.;  
 RT "Nine unique repeating sequences in a region essential for  
 RT replication and incompatibility of the mini-F plasmid."  
 RL Gene 15:257-271(1981).  
 CC (2)  
 RP SEQUENCE FROM N.A.  
 RX Eichenlaub R.;  
 RT "F Plasmid DNA complete mini-F region (F coordinates 40,301F to  
 RT 49,869F)." ;  
 RL Submitted (Aug-1986) to the EMBL/GenBank/DBJ databases.  
 RN [3]

RP SEQUENCE FROM N.A.  
RC STRAIN-K12 / CR63;  
RC Shimizu H., Saitoh Y., Suda Y., Uehara K., Sampel G., Mizobuchi K.;  
RT "Complete nucleotide sequence of the F plasmid: its implications for  
RT organization and diversification of plasmid genomes."  
RL Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.  
RN  
RP SEQUENCE OF 211-251 FROM N.A.  
RX MEDLINE-8134554; PubMed-7018695;  
RA Tolun A., Helinski D.R.;  
RT "Direct repeats of the F plasmid *lncC* region express F  
RT incompatibility."  
RL Cell 24:687-694(1981).  
RN  
RP SEQUENCE OF 1-17 FROM N.A.  
RX MEDLINE-86174346; PubMed-3007930;  
RA Disque-Kochem C., Seidel U., Helsenberg M., Eichenlaub R.;  
RT "The repeated sequences (*lncB*) preceding the protein E gene of  
RT plasmid *min-F* are essential for replication."  
RL Mol. Genet. 202:132-135(1986).  
RN  
RP SEQUENCE OF 209-251 FROM N.A.  
RC STRAIN-K12;  
RA MEDLINE-87141188; PubMed-3029390;  
RA Mori H., Kondo A., Onshima A., Ogura T., Hiraga S.;  
RT "Structure and function of the F plasmid genes essential for  
RT partitioning."  
RL J. Mol. Biol. 192:1-15(1986).  
CC -1- SIMILARITY: REGIONS OF SIMILARITY WITH REPA PROTEIN OF THE P1  
CC PLASMID, BUT NOT WITH REPLICATION PROTEINS OF OTHER PLASMIDS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
DR EMBL: J01724; AAA91062.1; -  
DR EMBL: M12987; AAA24901.1; -  
DR EMBL: A001918; BAA97915.1; -  
DR EMBL: X03410; CAA27146.1; -  
DR EMBL: X04619; CAA28294.1; -  
DR PIR: A04485; OQECF.  
DR PIR: D25783; D25783.  
DR Ecogene: EG40058; repE.  
DR InterPro: IPR000525; IntPr-Rep.  
DR Pfam: PF01051; RepB\_protein; 1.  
DR Plasmid: DNA replication; Plasmid copy control; Complete proteome.  
SQ SEQUENCE 251 AA; 29358 MW; EA2BC265D80158B2 CRC64;  
QY 1 KXXXTXQEXD 10  
DB 48 KSDGTQEHND 57  
Query Match 71.0%; Score 22; DB 1; Length 251;  
Best Local Similarity 50.0%; Pred. No. 36;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OX Streptococcus.  
OC NCBI\_TaxID=1314, 186103;  
RN  
RP SEQUENCE FROM N.A.  
RX MEDLINE-86166804; PubMed-3514452;  
RA Weeks C.R., Ferretti J.J.;  
RT "Nucleotide sequence of the type A streptococcal exotoxin  
RT (erythrogenic toxin) gene from Streptococcus pyogenes bacteriophage  
RT T12."  
RL Infect. Immun. 52:144-150(1986).  
RN  
RP SEQUENCE FROM N.A.  
RX MEDLINE-86284313; PubMed-3526093;  
RA Johnson L.P., L'Italian J.J., Schlievert P.M.;  
RT "Streptococcal pyrogenic exotoxin type A (scarlet fever toxin) is  
RT related to Staphylococcus aureus enterotoxin B."  
RL Mol. Genet. 203:354-356(1986).  
RN  
RP SEQUENCE FROM N.A.  
RC STRAIN-MGAS8232 / Serotype M18;  
RA MEDLINE-21927593; PubMed-11917108;  
RA Smoot J.C., Bardlan K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,  
RA Sylvia G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,  
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,  
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;  
RT "Genome sequence and comparative microarray analysis of serotype M18  
RT group A Streptococcus strains associated with acute rheumatic fever  
RT outbreaks."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).  
RN  
RN X-RAY CRYSTALLOGRAPHY (2.57 ANGSTROMS).  
RX MEDLINE-99094887; PubMed-9878045;  
RA Papageorgiou A.C., Collins C.M., Gutman D.M., Kline J.B.,  
RA O'Brien S.M., Tranter H.S., Acharya K.R.;  
RT "Structural basis for the recognition of superantigen streptococcal  
RT pyrogenic exotoxin A (SpeA) by MHC class II molecules and T-cell  
RT receptors."  
RL EMBO J. 18:9-21(1999).  
CC -1- SUBUNIT: Binds to major histocompatibility complex class II beta  
CC chain.  
CC -1- DISEASE: THE STREPTOCOCCAL PYROGENIC TOXINS A, B, AND C ARE  
CC THE CAUSATIVE AGENTS OF THE SYMPTOMS ASSOCIATED WITH SCARLET  
CC FEVER, HAVE BEEN ASSOCIATED WITH STREPTOCOCCAL TOXIC SHOCK-LIKE  
CC DISEASE AND MAY PLAY A ROLE IN THE EARLY EVENTS OF RHEUMATIC  
CC FEVER.  
CC -1- MISCELLANEOUS: THIS TOXIN SEEMS TO BE CODED BY BACTERIOPHAGE T12.  
CC -1- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN  
CC FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
DR EMBL: U40453; AAC48868.1; -  
DR EMBL: X03929; CAA27568.1; -  
DR EMBL: A000982; AAL97141.1; -  
DR PIR: A26152; A26152.  
DR PIR: S29659; S29659.  
DR PDB: 1B12; 24-NOV-99.  
DR InterPro: IPR001961; Staph/Strep\_toxin.  
DR Pfam: PF01123; Staph\_Strep\_toxin; 1.  
DR Pfam: PF02876; Staph\_Strep\_tox\_C; 1.  
DR PRINTS: PR00279; BACSTRTXIN.  
DR PROSITE: PS00277; STAPH\_STREP\_TOXIN\_1; 1.  
DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.  
KW Toxin; Signal; 3D-structure.  
FT SIGNAL 1 30  
FT CHAIN 31 251 EXOTOXIN TYPE A.

```

FT DISULFID 117 128 K -> E (IN REF. 2).
FT CONFLICT 6 6 VT -> MK (IN REF. 2).
FT CONFLICT 17 18 SOEFAOODPD -> LPFGICSTRPK (IN REF. 2).
FT CONFLICT 25 35 H -> O (IN REF. 2).
FT CONFLICT 40 40 S -> N (IN REF. 2).
FT CONFLICT 43 43 .NONTYFLIEGDP -> TEKTIYFMRVTL (IN REF. 2).
FT CONFLICT 47 59 I -> L (IN REF. 2).
FT CONFLICT 129 129 TNKKMTAQELDLYK -> QIKNGNCSRISTY (IN REF. 2).
FT CONFLICT 165 178
SQ SEQUENCE 251 AA; 29246 MM; 54001FEACCCBCC3 CRC64;

Query Match 71.0%; Score 22; DB 1; Length 251;
Best Local Similarity 50.0%; Pred. No. 36;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10
DB 167 KKKVTAQELD 176

RESULT 9
EXXB_STAAU STANDARD; PRT; 266 AA.
AC P01552;
DT 21-JUL-1986 (Rel. 01, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Enterotoxin type B precursor (SEB).
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=66168029; PubMed=3957869;
RA Jones C.L., Khan S.A.;
RT "Nucleotide sequence of the enterotoxin B gene from Staphylococcus aureus."
RL J. Bacteriol. 166:29-33(1986).
[2]
RP SEQUENCE OF 40-91 FROM N.A.
RX MEDLINE=85298255; PubMed=3698073;
RA Raneli D.M., Jones C.L., Johns M.B., Mussey G.J., Khan S.A.;
RT "Molecular cloning of staphylococcal enterotoxin B gene in Escherichia coli and Staphylococcus aureus."
RL Proc. Natl. Acad. Sci. U.S.A. 82:5850-5854(1985).
[3]
RP SEQUENCE OF 28-266 (S-6).
RX MEDLINE=71007902; PubMed=5470821;
RA Huang I.-Y., Bergdoll M.S.;
RT "Type primary structure of staphylococcal enterotoxin B. 3. The cyanogen bromide peptides of reduced and aminoethylated enterotoxin B, and the complete amino acid sequence."
RL J. Biol. Chem. 245:3518-3525(1970).
[4]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=93063291; PubMed=1436058;
RA Swaminathan S., Furey W.F., Jr., Pletcher J., Sax M.;
RT "Crystal structure of staphylococcal enterotoxin B, a superantigen."
RL Nature 359:801-806(1992).
[5]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF COMPLEX WITH MHC II.
RX MEDLINE=94203282; PubMed=8152483;
RA Janderzky T.S., Brown J.H., Gorga J.C., Stern L.J., Urban R.G., Chi Y.I., Stauffer C., Strominger J.L., Wiley D.C.;
RT "Three-dimensional structure of a human class II histocompatibility molecule complexed with superantigen."
RL Nature 368:711-718(1994).
[6]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF COMPLEX WITH TCR.
RX MEDLINE=99096298; PubMed=9681971;

```

```

RA Li H., Ilera A., Tsuchiya D., Leder L., Ysearn X., Schlievert P.M.,
RA Karjalainen K., Mariuzza R.A.;
RT "Three-dimensional structure of the complex between a T cell receptor beta chain and the superantigen staphylococcal enterotoxin B."
RL Immunity 9:807-816(1998).
[7]
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RX MEDLINE=98181012; PubMed=9514739;
RA Papaigeorgiou A.C., Tranter H.S., Acharya K.R.;
RT "Crystal structure of microbial superantigen staphylococcal enterotoxin B at 1.5-A resolution: implications for superantigen recognition by MHC class II molecules and T-cell receptors."
RL J. Mol. Biol. 277:61-79(1998).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION STAPHYLOCOCCAL FOOD POISONING SYNDROME.
CC -1- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M11118; AAA8550.1;
DR PIR: A01815; ENSAB6.
DR PIR: S27360; S27360.
DR PDB: 1SEB; 20-JUN-96.
DR PDB: 2SEB; 28-JAN-98.
DR PDB: 3SEB; 27-MAY-98.
DR PDB: 1SE3; 16-JUN-97.
DR PDB: 1SE4; 15-OCT-97.
DR PDB: 1SEB; 04-MAR-99.
DR InterPro: IPR001961; Staph/strep_toxin.
DR Pfam: PF01123; Staph_strep_toxin; 1.
DR Pfam: PF02876; Staph_strep_toxin; 1.
DR PRINTS: PR00279; BACRLOTOXIN.
DR PROSITE: PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE: PS00278; STAPH_STREP_TOXIN_2; 1.
DR Enterotoxin; Toxin; Signal; Superantigen; 3D-structure.
KW SIGNAL
FT SIGNAL 1 27
FT DISULFID 28 266 ENTEROTOXIN TYPE B.
FT CONFLICT 56 58
FT CONFLICT 69 77 DDN -> NND (IN REF. 3).
FT CONFLICT 118 118 DOFLIEDL -> NEFDLITL (IN REF. 3).
FT CONFLICT 128 130 DIN -> NID (IN REF. 3).
FT CONFLICT 133 135 QTD -> ENT (IN REF. 3).
FT CONFLICT 149 150 NG -> GN (IN REF. 3).
FT CONFLICT 156 156 Y -> YY (IN REF. 3).
FT CONFLICT 185 186 OE -> EO (IN REF. 3).
FT CONFLICT 233 233 D -> N (IN REF. 3).
FT CONFLICT 246 247 DN -> ND (IN REF. 3).
SQ SEQUENCE 266 AA; 31436 MM; B6D417F61CF018B0 CRC64;

Query Match 71.0%; Score 22; DB 1; Length 266;
Best Local Similarity 50.0%; Pred. No. 39;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10
DB 179 KKKVTAQELD 188

RESULT 10
G3P3_YEAST STANDARD; PRT; 331 AA.
AC P00359;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)

```

DT 15-UTN-2002 (rel. 41, Last annotation update)  
 DE Glyceraldehyde 3-phosphate dehydrogenase 3 (EC 1.2.1.12) (GAPDH 3).  
 GN TDH3 OR GPD3 OR YGR192C OR G7576.  
 OS Saccharomyces cerevisiae (Baker's Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=80027306; PubMed=385592;  
 RA Holland J.P., Holland M.J.;  
 RT "The primary structure of a glyceraldehyde-3-phosphate dehydrogenase  
 RT gene from Saccharomyces cerevisiae.";  
 RL J. Biol. Chem. 254:9839-9845(1979).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=S288C;  
 RA MEDLINE=95373283; PubMed=7645350;  
 RA Arroyo J., Garcia-Gonzalez M., Garcia-Saez M.I., Sanchez M.,  
 RA Nombela C.;  
 RT "The complete sequence of a 9037 bp DNA fragment of the right arm of  
 RT Saccharomyces cerevisiae chromosome VII.";  
 RL Yeast 11:587-591(1995).  
 RN [3]  
 RP SEQUENCE OF 46-57.  
 RX STRAIN=S288C;  
 RA MEDLINE=95203288; PubMed=7895733;  
 RA Garrels J.I., Futcher B., Kobayashi R., Latter G.I., Schwender B.,  
 RA Volpe T., Warner J.R., McLaughlin C.S.;  
 RT "Protein identifications for a Saccharomyces cerevisiae protein  
 RT database.";  
 RL Electrophoresis 15:1466-1486(1994).  
 RN [4]  
 RP PARTIAL SEQUENCE.  
 RX STRAIN=ATCC 38531 / Y41, and SKO2N;  
 RX MEDLINE=95555188; PubMed=7737086;  
 RA Norbeck J., Blomberg A.;  
 RT "Gene linkage of two-dimensional polyacrylamide gel electrophoresis  
 RT resolved proteins from isogene families in Saccharomyces cerevisiae  
 RT by microsequencing of in-gel trypsin generated peptides.";  
 RL Electrophoresis 16:149-156(1995).  
 CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +  
 CC NAD(+) -> 3-phospho-D-glyceroyl phosphate + NADH.  
 CC -1- PATHWAY: Second phase of glycolysis; first step.  
 CC -1- SUBUNIT: HOMOTETRAMER.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- MISCELLANEOUS: THERE ARE THREE GENES FOR GAPDH IN YEAST.  
 CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE  
 CC DEHYDROGENASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: V01300; CA24607.1; -  
 DR EMBL: J01324; AA88714.1; -  
 DR EMBL: X82408; CA57803.1; -  
 DR EMBL: Z72977; CA97218.1; -  
 DR PIR: A00371; DEBYG2.  
 DR HSSP: P06977; IGAD.  
 DR HSSP-2DPAGE: P00359; YEAST.  
 DR COMPUYEST-2DPAGE: P00359; -  
 DR SGD: S0003424; TDH3.  
 DR InterPro: IPR000173; GAP\_dhndrogenase.  
 DR Pfam: PF00044; gpdh.1.  
 DR PRINTS: PR00078; G3PDHRCNASE.  
 DR PROSITE: PS00071; GAPDH.1.  
 KW Glycolysis; Oxidoreductase; NAD; Multigene family.

FT INT MET 0 0  
 FT BINDING 149 149  
 FT ACT SITE 176 176  
 FT CONFLICT 135 135  
 FT CONFLICT 247 247  
 FT CONFLICT 328 328  
 SQ SEQUENCE 331 AA; 35615 MW; CFPE94A335C648B5 CRC64;  
 Query Match 71.0%; Score 22; DB 1; Length 331;  
 Best Local Similarity 50.0%; Pred. No. 49;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 OY 1 KXXXTXQEND 10  
 DB 69 KKIATYQEND 78  
 RESULT 11  
 DRK3.CHICK STANDARD; PRT; 350 AA.  
 AC 090839;  
 DT 01-NOV-1997 (rel. 35, Created)  
 DT 01-NOV-1997 (rel. 35, Last sequence update)  
 DT 16-OCT-2001 (rel. 40, Last annotation update)  
 DE Dickkopf related protein-3 precursor (Dkk-3) (Dickkopf-3) (Lens fiber  
 DE protein CLFE5T4).  
 GN DRK3.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archaeopteryx; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE= lens fibers;  
 RX MEDLINE=96437509; PubMed=8840185;  
 RA Sawada K., Agata K., Eguchi G.;  
 RT "Characterization of terminally differentiated cell state by  
 RT categorizing cDNA clones derived from chicken lens fibers.";  
 RL Int. J. Dev. Biol. 40:531-535(1996).  
 CC -1- FUNCTION: INHIBITOR OF WNT SIGNALING PATHWAY (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN EYE LENS.  
 CC -1- SIMILARITY: BELONGS TO THE DICKKOPF FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: D26311; BAA05373.1; -  
 DR HSSP: P25687; IIMT.  
 KW Developmental protein; Signal; Glycoprotein.  
 FT SIGNAL 1 29  
 FT CHAIN 1 350  
 FT DOMAIN 139 187  
 FT DOMAIN 200 277  
 FT CARBOHYD 88 88  
 FT CARBOHYD 98 98  
 FT CARBOHYD 113 113  
 FT CARBOHYD 196 196  
 FT CARBOHYD 282 282  
 SQ SEQUENCE 350 AA; 39208 MW; 57BE7ED850089DAE CRC64;  
 Query Match 71.0%; Score 22; DB 1; Length 350;  
 Best Local Similarity 50.0%; Pred. No. 52;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 OY 1 KXXXTXQEND 10

DB 99 KTVOTHOEID 108

RESULT 12  
YDGC\_SCHPO STANDARD: PRT: 356 AA.

ID YDGC\_SCHPO STANDARD: PRT: 356 AA.

AC 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Hypothetical protein C26F1.12c in chromosome I.

GN SPAC26F1.12C

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI\_TaxID=4896;

RT [1]

RP SEQUENCE FROM N.A.

RC STRAIN=972;

RA MEDLINE=21848401; PubMed=11859360;

RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,

RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,

RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,

RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,

RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,

RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,

RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,

RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,

RA Skelton J., Stimmings M., Squares R., Squares S., Stevens K.,

RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,

RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,

RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,

RA Gabel C., Fuchs M., Fritz C., Holzer E., Moesli D., Hilbert H.,

RA Borzym K., Langer I., Beck A., Lehmach H., Reinhardt R., Pohl T.M.,

RA Egger P., Zimmermann W., Wedler H., Wandut R., Purnelle B.,

RA Goffeau A., Cadiou E., Dreano S., Gloux S., Lelaure V., Mottier S.,

RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Rhode G.,

RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,

RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,

RA Cerruti L., Love T., McCombie W.R., Paulsen I., Potashkin J.,

RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;

RT "The genome sequence of *Schizosaccharomyces pombe*."

RL Nature 415:871-880 (2002).

CC -1- SIMILARITY: TO YEAST HGHL.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

DR EMBL: 273100; CAA97369.1; -

DR Hypothetical protein.

DR KMW

DR SEQUENCE 356 AA; 41260 MW; AD1BC714C134AA1C CRC64;

Query Match 71.0%; Score 22; DB 1; Length 356;

Best Local Similarity 50.0%; Pred. No. 53;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10

DB 178 KYATNGEYD 187

RESULT 13

YB55\_RICPR STANDARD: PRT: 520 AA.

ID YB55\_RICPR STANDARD: PRT: 520 AA.

AC Q9ZCA5;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein RP855.

GN RP855.

OS Rickettsia prowazekii.

OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;

OC Rickettsiaceae; Rickettsiae; Rickettsia.

OX NCBI\_TaxID=782;

RT [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Madrid E;

RA MEDLINE=99039499; PubMed=9823893;

RA Anderson S.G.E., Zomorodipour A., Andersson J.O.,

RA Sichteritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,

RA Eriksson A.-S., Winkler H.H., Kurland C.G.;

RT "The genome sequence of *Rickettsia prowazekii* and the origin of

RT mitochondria."

RL Nature 396:133-140 (1998).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

DR EMBL: AJ235273; CAA15279.1; -

DR Hypothetical protein; Transmembrane; Complete proteome.

DR KW

DR TRANSMEM 337

DR SEQUENCE 520 AA; 61113 MW; A4B4DBA1EE594034 CRC64;

Query Match 71.0%; Score 22; DB 1; Length 520;

Best Local Similarity 50.0%; Pred. No. 78;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10

DB 374 KYATNGEYD 383

RESULT 14

ARS2\_CAEEL STANDARD: PRT: 712 AA.

ID ARS2\_CAEEL STANDARD: PRT: 712 AA.

AC 0966L5;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Arsenite-resistance protein 2 homolog.

GN E01A2.2.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

OC Rhabditidae; Pelodertinae; Caenorhabditis.

OX NCBI\_TaxID=6239;

RT [1]

RP SEQUENCE FROM N.A.

RC STRAIN-Bristol N2;

RA Wilson R., Bentley D., Scheet P.;

RA Submitted (Mar-1999) to the EMBL/GenBank/DBJ databases.

RL -1- SIMILARITY: BELONGS TO THE ARS2 FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

DR EMBL: AC006627; AAK85459.1; -

DR WormPep; E01A2.2; CE20643.



KW Hypothetical protein.  
SQ SEQUENCE 712 AA; 80794 MW; 54B051ADFA547DE CRC64;

Query Match 71.0%; Score 22; DB 1; Length 712;  
Best Local Similarity 50.0%; Pred. No. 1.4e+02;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXXXXQEXD 10  
| | | | |  
Db 88 KKFLLTSQEDD 97

RESULT 15  
YE86\_SCHPO  
ID YE86\_SCHPO STANDARD; PRT; 886 AA.  
AC 014302;

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Hypothetical protein C9G1.06c in chromosome I.

GN SPAC9G1.06c.

OS Schizosaccharomyces pombe (fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

NCBI\_TaxID=4896;

RP SEQUENCE FROM N.A.

RC STRAIN=972:

RA MEDLINE=21848401; PubMed=11859360;

RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,

RA Brooks K., Brown D., Brown S., Chillingworth T., Church C.M.,

RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,

RA Holtroyd S., Hornsby I., Howarth S., Huckle E.J., Hunt S., Jagsels K.,

RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,

RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabblnowitsch E.,

RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,

RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,

RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,

RA Woodward J., Volckaert G., Aert R., Robben J., Gymnopoulos B.,

RA Wellens J., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,

RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,

RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,

RA Eger P., Zimmermann W., Medler H., Wandt R., Purnelle B.,

RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,

RA Gallbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,

RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,

RA Dominguez A., Revelante J.L., Moreno S., Armstrong J., Forsburg S.L.,

RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,

RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.,

RT The genome sequence of Schizosaccharomyces pombe.\*;

RL Nature 415:871-880(2002).

CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.

CC -1- SIMILARITY: TO YEAST YDL117W.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: Z98763; CAB11490.1; -

DR HSSP: P29355; ISEM.

DR InterPro: IPR001452; SH3.

DR InterPro: IPR002931; Trnsglucanase\_like.

DR Pfam: PF000018; SH3; 1.

DR ProDom: PD000066; SH3; 1.

DR SMART: SM00326; SH3; 1.

DR SMART: SM00460; TGC; 1.

DR PROSITE: PS50002; SH3; 1.

KW Hypothetical protein: SH3 domain.

FT DOMAIN 6 67 SH3.

FT POLY-SER. 607 610

SQ SEQUENCE 886 AA; 98262 MW; D0C4879C1882E869 CRC64;

Query Match 71.0%; Score 22; DB 1; Length 886;  
Best Local Similarity 40.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXXXXQEXD 10  
| | | | |  
Db 220 KSTTSKEDD 229

Search completed: June 5, 2003, 15:56:31

Job time : 8.51515 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw'model

Run on: June 5, 2003, 15:32:58 ; Search time 29.0909 Seconds

(without alignments)  
70.829 Million cell updates/sec

Title: US-09-150-947f-13

Perfect score: 31

Sequence: 1 KXXXTXQEXD 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	24	77.4	209	2	09R005	09R95 streptococc
2	24	77.4	209	2	09LAE1	09LAE1 streptococc
3	24	77.4	209	2	09LAE0	09LAE0 streptococc
4	24	77.4	209	2	09LAD9	09LAD9 streptococc
5	24	77.4	209	2	09LAD8	09LAD8 streptococc
6	24	77.4	209	2	09LAD7	09LAD7 streptococc
7	24	77.4	209	2	09LAD6	09LAD6 streptococc
8	24	77.4	209	2	09LAD5	09LAD5 streptococc
9	24	77.4	209	2	09LAD4	09LAD4 streptococc
10	24	77.4	209	2	09LAD3	09LAD3 streptococc
11	24	77.4	209	2	09LAD2	09LAD2 streptococc
12	24	77.4	209	2	09LAD1	09LAD1 streptococc
13	24	77.4	209	2	09LAD0	09LAD0 streptococc
14	24	77.4	209	2	09LAC9	09LAC9 streptococc
15	24	77.4	209	2	09LAC8	09LAC8 streptococc
16	24	77.4	209	2	09LAC7	09LAC7 streptococc

17	24	77.4	209	2	09LAC6	09LAC6 streptococc
18	24	77.4	209	2	09LAC5	09LAC5 streptococc
19	24	77.4	209	2	09LAC4	09LAC4 streptococc
20	24	77.4	209	2	09LAC3	09LAC3 streptococc
21	24	77.4	233	16	099XW1	099XW1 streptococc
22	24	77.4	256	2	09XSR8	09XSR8 streptococc
23	24	77.4	256	2	09S1H9	09S1H9 streptococc
24	24	77.4	256	2	09S1H8	09S1H8 streptococc
25	24	77.4	549	10	09LND29	09LND29 arabidopsis
26	24	77.4	603	12	084742	084742 human para
27	24	77.4	603	12	081076	081076 human para
28	24	77.4	955	11	090Y37	090Y37 mus musculu
29	24	77.4	955	11	08VHL8	08VHL8 mus musculu
30	23	74.2	234	2	09R5X4	09R5X4 staphylococ
31	23	74.2	239	2	053678	053678 staphylococ
32	23	74.2	239	2	005157	005157 staphylococ
33	23	74.2	239	2	006531	006531 staphylococ
34	23	74.2	239	2	006532	006532 staphylococ
35	23	74.2	239	2	006533	006533 staphylococ
36	23	74.2	239	2	006534	006534 staphylococ
37	23	74.2	239	2	006535	006535 staphylococ
38	23	74.2	271	2	09F0L6	09F0L6 staphylococ
39	23	74.2	324	4	043370	043370 homo sapien
40	23	74.2	405	2	09XBU4	09XBU4 bacillus ce
41	23	74.2	405	2	08VW49	08VW49 bacillus th
42	23	74.2	561	2	09X4E1	09X4E1 rhodobacter
43	23	74.2	565	16	09A2H4	09A2H4 caulobacter
44	23	74.2	565	16	098CC3	098CC3 rhizobium l
45	23	74.2	572	16	080818	080818 agrobacteri

#### ALIGNMENTS

##### RESULT 1

ID 09R005 PRELIMINARY: PRT: 209 AA.  
AC 09R005;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAY-2002 (TREMBLrel. 20, Last annotation update)  
DE Mitogenic exotoxin Z 2 (Fragment).  
GN SMEZ-2.  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
OC Streptococcaceae; Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=2035;  
RX MEDLINE=99093428; PubMed=9874566;  
RA Prof. T. Woffatt S.L., Berkehn C.J., Fraser J.D.;  
RT "Identification and characterization of novel superantigens from  
RT Streptococcus pyogenes."  
RL J. Exp. Med. 189:89-102(1999).  
DR EMBL: AF086626; AAD52087.1; -.  
DR HSSP: P13163; ISXT.  
DR InterPro: IPR001961; Staph/Strep toxin.  
DR Pfam: PF01123; Staph\_strep\_toxin\_1.  
DR PRINTS: PR00279; BACTRLTOXIN.  
DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.  
FT NON\_TER  
SQ SEQUENCE 209 AA; 24131 MW; 52BF7911BB100152 CRC64;

Query Match 1 KXXXTXQEXD 10 77.4% Score 24; DB 2; Length 209;  
Best Local Similarity 50.0%; Pred. No. 46;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

DB 125 KTVVTAQED 134

```

RESULT 2
Q9LAE1          PRELIMINARY;      PRT;      209 AA.
AC Q9LAE1:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Mitogenic exotoxin 2-3 (Fragment).
GN SMEZ-3.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=11681;
RX MEDLINE=20273982; PubMed=10811869;
RA Proft T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,
  Fraser J.D.;
RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,
  Mosaic Structure, and Significant Antigenic Variation."
RL J. Exp. Med. 191:1765-1776(2000).
DR HSSP; P13163; 1SXT.
DR InterPro; IPR001961; Staph/Strep-toxin.
DR Pfam; PF01123; Staph_Strep_toxin; 1.
DR PRINTS; PR002876; Staph_Strep_toxin; 1.
DR PROSITE; PS00279; BACTRLTOXIN.
DR NON_TER 1
SQ SEQUENCE 209 AA; 24071 MW; FPAEDCDIAA87271 CRC64;

Query Match
Best Local Similarity 77.4%; Score 24; DB 2; Length 209;
Matches 5; Conservative 50.0%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10
Db 125 KTYVTAQED 134

RESULT 3
Q9LAE0          PRELIMINARY;      PRT;      209 AA.
AC Q9LAE0:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Mitogenic exotoxin 2-4 (Fragment).
GN SMEZ-4.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9893;
RX MEDLINE=20273982; PubMed=10811869;
RA Proft T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,
  Fraser J.D.;
RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,
  Mosaic Structure, and Significant Antigenic Variation."
RL J. Exp. Med. 191:1765-1776(2000).
DR HSSP; P13163; 1SXT.
DR InterPro; IPR001961; Staph/Strep-toxin.
DR Pfam; PF01123; Staph_Strep_toxin; 1.
DR PRINTS; PR002876; Staph_Strep_toxin; 1.
DR PROSITE; PS00279; BACTRLTOXIN.
DR NON_TER 1
SQ SEQUENCE 209 AA; 24108 MW; 67EC279BCC4A8247 CRC64;

```

```

Query Match
Best Local Similarity 77.4%; Score 24; DB 2; Length 209;
Matches 5; Conservative 50.0%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10
Db 125 KTYVTAQED 134

RESULT 4
Q9LAD9          PRELIMINARY;      PRT;      209 AA.
AC Q9LAD9:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Mitogenic exotoxin 2-5 (Fragment).
GN SMEZ-5.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=11244;
RX MEDLINE=20273982; PubMed=10811869;
RA Proft T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,
  Fraser J.D.;
RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,
  Mosaic Structure, and Significant Antigenic Variation."
RL J. Exp. Med. 191:1765-1776(2000).
DR HSSP; P13163; 1SXT.
DR InterPro; IPR001961; Staph/Strep-toxin.
DR Pfam; PF01123; Staph_Strep_toxin; 1.
DR PRINTS; PR002876; Staph_Strep_toxin; 1.
DR PROSITE; PS00279; BACTRLTOXIN.
DR NON_TER 1
SQ SEQUENCE 209 AA; 24079 MW; 24CA3885469CBB9B CRC64;

Query Match
Best Local Similarity 77.4%; Score 24; DB 2; Length 209;
Matches 5; Conservative 50.0%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10
Db 125 KTYVTAQED 134

RESULT 5
Q9LAD8          PRELIMINARY;      PRT;      209 AA.
AC Q9LAD8:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Mitogenic exotoxin 2-7 (Fragment).
GN SMEZ-7.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=11574;
RX MEDLINE=20273982; PubMed=10811869;
RA Proft T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,
  Fraser J.D.;
RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,
  Mosaic Structure, and Significant Antigenic Variation."
RL J. Exp. Med. 191:1765-1776(2000).

```

DR EMBL: AF143657; AAF66657.1; -.  
 DR HSSP; P13163; 1SXT.  
 DR InterPro: IPR001961; Staph/Strep\_toxin.  
 DR Pfam: PF01123; Staph\_Strep\_toxin; 1.  
 DR Pfam: PF02876; Staph\_Strep\_tox\_C; 1.  
 DR PRINTS: PR00279; BACTRLTOXIN.  
 DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 209 AA; 24064 MW; 67B977CD5AA934F7 CRC64;

Query Match 77.4%; Score 24; DB 2; Length 209;  
 Best Local Similarity 50.0%; Pred. No. 46;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
 | | | | |  
 DB 125 KTTVTAQEXD 134

## RESULT 6

ID 09LAD7 PRELIMINARY; PRT; 209 AA.  
 AC 09LAD7:  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Mitogenic exotoxin Z-8 (Fragment).  
 GN SMEZ-8.  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
 OC Streptococcaceae; Streptococcus.  
 OX NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1/5045;  
 RX MEDLINE-20273982; PubMed-10811869;  
 RA Proft T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,  
 RA Frazer J.D.;  
 RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,  
 RT Mosaic Structure, and Significant Antigenic Variation.";  
 RL J. Exp. Med. 191:1765-1776(2000).  
 DR EMBL: AF143658; AAF66658.1; -.  
 DR HSSP; P13163; 1SXT.  
 DR InterPro: IPR001961; Staph/Strep\_toxin.  
 DR Pfam: PF01123; Staph\_Strep\_toxin; 1.  
 DR Pfam: PF02876; Staph\_Strep\_tox\_C; 1.  
 DR PRINTS: PR00279; BACTRLTOXIN.  
 DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 209 AA; 24091 MW; 7680E9B3513DF28 CRC64;

Query Match 77.4%; Score 24; DB 2; Length 209;  
 Best Local Similarity 50.0%; Pred. No. 46;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
 | | | | |  
 DB 125 KTTVTAQEXD 134

## RESULT 7

ID 09LAD6 PRELIMINARY; PRT; 209 AA.  
 AC 09LAD6:  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Mitogenic exotoxin Z-9 (Fragment).  
 GN SMEZ-9.  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
 OC Streptococcaceae; Streptococcus.  
 OX NCBI\_TaxID=1314;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=11299;  
 RX MEDLINE-20273982; PubMed-10811869;  
 RA Proft T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,  
 RA Frazer J.D.;  
 RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,  
 RT Mosaic Structure, and Significant Antigenic Variation.";  
 RL J. Exp. Med. 191:1765-1776(2000).  
 DR EMBL: AF143659; AAF66659.1; -.  
 DR HSSP; P13163; 1SXT.  
 DR InterPro: IPR001961; Staph/Strep\_toxin.  
 DR Pfam: PF01123; Staph\_Strep\_toxin; 1.  
 DR Pfam: PF02876; Staph\_Strep\_tox\_C; 1.  
 DR PRINTS: PR00279; BACTRLTOXIN.  
 DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 209 AA; 24213 MW; B21587BA33B5DD5 CRC64;

Query Match 77.4%; Score 24; DB 2; Length 209;  
 Best Local Similarity 50.0%; Pred. No. 46;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
 | | | | |  
 DB 125 KTTVTAQEXD 134

## RESULT 8

ID 09LAD5 PRELIMINARY; PRT; 209 AA.  
 AC 09LAD5:  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Mitogenic exotoxin Z-10 (Fragment).  
 GN SMEZ-10.  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
 OC Streptococcaceae; Streptococcus.  
 OX NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=11580;  
 RX MEDLINE-20273982; PubMed-10811869;  
 RA Proft T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,  
 RA Frazer J.D.;  
 RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,  
 RT Mosaic Structure, and Significant Antigenic Variation.";  
 RL J. Exp. Med. 191:1765-1776(2000).  
 DR EMBL: AF143660; AAF66660.1; -.  
 DR HSSP; P13163; 1SXT.  
 DR InterPro: IPR001961; Staph/Strep\_toxin.  
 DR Pfam: PF01123; Staph\_Strep\_toxin; 1.  
 DR Pfam: PF02876; Staph\_Strep\_tox\_C; 1.  
 DR PRINTS: PR00279; BACTRLTOXIN.  
 DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 209 AA; 24102 MW; 8B4C853751A56D06 CRC64;

Query Match 77.4%; Score 24; DB 2; Length 209;  
 Best Local Similarity 50.0%; Pred. No. 46;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
 | | | | |  
 DB 125 KTTVTAQEXD 134

## RESULT 9

ID 09LAD4 PRELIMINARY; PRT; 209 AA.  
 AC 09LAD4:

```

DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Mitogenic exotoxin 2-11 (Fragment).
GN SMEZ-11.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9779;
RX MEDLINE=20273982; PubMed=10811869;
RA Proft T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,
  Fraser J.D.;
RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,
  Mosaic Structure, and Significant Antigenic Variation.";
RL J. Exp. Med. 191:1765-1776(2000).
DR EMBL; AF143661; AAF66661.1; -.
DR HSSP; P13163; 1SXT.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_Strep_toxin; 1.
DR Pfam; PF02876; Staph_Strep_toxin_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
FT NON_TER 1
SQ SEQUENCE 209 AA; 24194 MW; E12EF47B38BD95DE CRC64;

Query Match 77.4%; Score 24; DB 2; Length 209;
Best Local Similarity 50.0%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXOEXD 10
DB 125 KTVTAQEI 134

RESULT 10
Q9LAD3 PRELIMINARY; PRT; 209 AA.
AC Q9LAD3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Mitogenic exotoxin 2-12 (Fragment).
GN SMEZ-12.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=10303;
RX MEDLINE=20273982; PubMed=10811869;
RA Proft T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,
  Fraser J.D.;
RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,
  Mosaic Structure, and Significant Antigenic Variation.";
RL J. Exp. Med. 191:1765-1776(2000).
DR EMBL; AF143662; AAF66662.1; -.
DR HSSP; P13163; 1SXT.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_Strep_toxin; 1.
DR Pfam; PF02876; Staph_Strep_toxin_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
FT NON_TER 1
SQ SEQUENCE 209 AA; 24086 MW; 1C4343235157DB68 CRC64;

Query Match 77.4%; Score 24; DB 2; Length 209;
Best Local Similarity 50.0%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 KXXXTXOEXD 10
DB 125 KTVTAQEI 134

RESULT 11
Q9LAD2 PRELIMINARY; PRT; 209 AA.
AC Q9LAD2;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Mitogenic exotoxin 2-13 (Fragment).
GN SMEZ-13.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=95/31;
RX MEDLINE=20273982; PubMed=10811869;
RA Proft T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,
  Fraser J.D.;
RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,
  Mosaic Structure, and Significant Antigenic Variation.";
RL J. Exp. Med. 191:1765-1776(2000).
DR EMBL; AF143663; AAF66663.1; -.
DR HSSP; P13163; 1SXT.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_Strep_toxin; 1.
DR Pfam; PF02876; Staph_Strep_toxin_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
FT NON_TER 1
SQ SEQUENCE 209 AA; 24072 MW; 984C1B4614589A1E CRC64;

Query Match 77.4%; Score 24; DB 2; Length 209;
Best Local Similarity 50.0%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXOEXD 10
DB 125 KTVTAQEI 134

RESULT 12
Q9LAD1 PRELIMINARY; PRT; 209 AA.
AC Q9LAD1;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Mitogenic exotoxin 2-14 (Fragment).
GN SMEZ-14.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=4202;
RX MEDLINE=20273982; PubMed=10811869;
RA Proft T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,
  Fraser J.D.;
RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,
  Mosaic Structure, and Significant Antigenic Variation.";
RL J. Exp. Med. 191:1765-1776(2000).
DR EMBL; AF143664; AAF66664.1; -.
DR HSSP; P13163; 1SXT.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_Strep_toxin; 1.
DR Pfam; PF02876; Staph_Strep_toxin_C; 1.

```

DR PRINTS: PR00279; BACTRLTOXIN.  
 DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.  
 RT NON\_TER 1  
 SQ SEQUENCE 209 AA; 24122 MW; 0CF5DA29E1B96EFE CRC64;

Query Match  
 Best Local Similarity 77.4%; Score 24; DB 2; Length 209;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
 DB 125 KTTVTAQED 134

RESULT 13  
 ID 09LAD0 PRELIMINARY; PRT; 209 AA.

AC 09LAD0: 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Mitogenic exotoxin Z-15 (Fragment).  
 GN SMEZ-15.

OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
 OC Streptococcaceae; Streptococcus.  
 OX NCBI\_Taxid=1314;

RP SEQUENCE FROM N.A.

RC STRAIN=;

RA MEDLINE-20273982; PubMed-10811869;

RA Profit T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,  
 RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,  
 RT Mosaic Structure, and Significant Antigenic Variation.";

RL J. Exp. Med. 191:1765-1776(2000).

DR EMBL: AF143665; AAF6665.1; -.

DR HSSP: P13163; 1SXT.

DR InterPro: IPR001961; Staph/Strep\_toxin.

DR Pfam: PF01123; Staph\_Strep\_toxin; 1.

DR Pfam: PF02876; Staph\_Strep\_toxin; 1.

DR PRINTS: PR00279; BACTRLTOXIN.

DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.

FT NON\_TER 1  
 SQ SEQUENCE 209 AA; 24209 MW; B909784ABCD1474 CRC64;

Query Match  
 Best Local Similarity 77.4%; Score 24; DB 2; Length 209;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
 DB 125 KTTVTAQED 134

RESULT 14  
 ID 09LAC9 PRELIMINARY; PRT; 209 AA.

AC 09LAC9: 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Mitogenic exotoxin Z-16 (Fragment).  
 GN SMEZ-16.

OS Streptococcus pyogenes.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
 OC Streptococcaceae; Streptococcus.

OX NCBI\_Taxid=1314;

RP SEQUENCE FROM N.A.

RC STRAIN=10649;

RX MEDLINE-20273982; PubMed-10811869;

RA Profit T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,

RA Fraser J.D.;  
 RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,  
 RT Mosaic Structure, and Significant Antigenic Variation.";  
 RL J. Exp. Med. 191:1765-1776(2000).

DR EMBL: AF143666; AAF6666.1; -.

DR HSSP: P13163; 1SXT.

DR InterPro: IPR001961; Staph/Strep\_toxin.

DR Pfam: PF01123; Staph\_Strep\_toxin; 1.

DR Pfam: PF02876; Staph\_Strep\_toxin; 1.

DR PRINTS: PR00279; BACTRLTOXIN.

DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.

FT NON\_TER 1  
 SQ SEQUENCE 209 AA; 24172 MW; A1DB8FA18709BEA5 CRC64;

Query Match  
 Best Local Similarity 77.4%; Score 24; DB 2; Length 209;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
 DB 125 KTTVTAQED 134

RESULT 15  
 ID 09LAC8 PRELIMINARY; PRT; 209 AA.

AC 09LAC8: 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Mitogenic exotoxin Z-17 (Fragment).

GN SMEZ-17.

OS Streptococcus pyogenes.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
 OC Streptococcaceae; Streptococcus.

OX NCBI\_Taxid=1314;

RP SEQUENCE FROM N.A.

RC STRAIN=11686;

RX MEDLINE-20273982; PubMed-10811869;

RA Profit T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,  
 RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,  
 RT Mosaic Structure, and Significant Antigenic Variation.";

RL J. Exp. Med. 191:1765-1776(2000).

DR EMBL: AF143667; AAF6667.1; -.

DR HSSP: P13163; 1SXT.

DR InterPro: IPR001961; Staph/Strep\_toxin.

DR Pfam: PF01123; Staph\_Strep\_toxin; 1.

DR Pfam: PF02876; Staph\_Strep\_toxin; 1.

DR PRINTS: PR00279; BACTRLTOXIN.

DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.

FT NON\_TER 1  
 SQ SEQUENCE 209 AA; 24181 MW; A333FA66398D9DC2 CRC64;

Query Match  
 Best Local Similarity 77.4%; Score 24; DB 2; Length 209;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
 DB 125 KTTVTAQED 134

Search completed: June 5, 2003, 15:55:40  
 Job time: 31.0909 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:47:43 ; Search time 33.9394 Seconds  
(without alignments)  
39,261 Million cell updates/sec

Title: US-09-150-947F-14

Perfect score: 26

Sequence: 1 KKKXXXXXLD 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*
- 19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*
- 20: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*
- 21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*
- 22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*
- 23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	73.1	378	21	Arabidopsis thaliana
2	18	69.2	46	21	Gene 3 human secre
3	18	69.2	46	21	Human secreted pro
4	18	69.2	147	21	Human secreted pro
5	18	69.2	152	19	Rat ninjurin 1. R
6	18	69.2	173	22	Human ninjurin 1.
7	18	69.2	173	22	Human secreted pro
8	18	69.2	200	19	Human secreted pro
9	18	69.2	213	22	Plant D-like cycli
10	18	69.2	247	13	pcTD ORF 6. Chlam

11	18	69.2	284	23	ABP41256
12	18	69.2	420	17	AAW93144
13	18	69.2	406	17	AAW93155
14	18	69.2	420	17	AAW93157
15	18	69.2	618	22	ABG19070
16	18	69.2	705	17	AAW06831
17	18	69.2	765	22	AAW73637
18	18	69.2	765	22	AAW94373
19	18	69.2	768	21	AAW42448
20	17	65.4	17	23	ABW76254
21	17	65.4	17	23	ABW76259
22	17	65.4	24	21	AAV97855
23	17	65.4	24	21	AAV97862
24	17	65.4	24	22	AAV72200
25	17	65.4	24	22	AAV72200
26	17	65.4	75	23	ABP01309
27	17	65.4	76	22	ABG15641
28	17	65.4	88	22	AAO11063
29	17	65.4	99	21	AAW35163
30	17	65.4	100	21	ABP32693
31	17	65.4	106	23	ABP32548
32	17	65.4	142	19	AAW48373
33	17	65.4	142	20	AAV28249
34	17	65.4	142	21	AAW01389
35	17	65.4	142	21	AAV39026
36	17	65.4	142	22	ABW5735
37	17	65.4	143	22	ABW12218
38	17	65.4	150	20	AAV36961
39	17	65.4	153	21	AAW60296
40	17	65.4	170	23	AAW90300
41	17	65.4	171	21	AAW35162
42	17	65.4	184	22	ABW66316
43	17	65.4	186	21	AAW42925
44	17	65.4	188	21	AAW60286
45	17	65.4	189	21	AAW60295

ALIGNMENTS

RESULT 1	
AAW35560	AAW35560 standard; Protein: 378 AA.
XX	
XX	AAW35560;
AC	
XX	
DT	18-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 43459.
XX	
XX	
KW	Protein identification: signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
PD	
XX	
XX	06-SEP-2000.
PF	
XX	25-FEB-2000; 2000EP-0301439.
XX	
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129645.
PR	19-APR-1999; 99US-0130077.

Human ovarian anti  
S. lividans xylana  
Streptomyces livid  
Streptomyces livid  
Novel human diapo  
Major neutralising  
Human ATP-dependen  
Human protein sequ  
Human ORF ORF2212  
Staphylococcus aur  
Staphylococcus aur  
Staphylococcus aur  
Peptide from regio  
Human ORF protein  
Novel human diapo  
Human polypeptide  
Zea mays protein f  
Zea mays protein f  
Human synthase-lik  
Human ninjurin 2.  
Human secreted pro  
Human secreted pro  
Human secreted pro  
Human polypeptide  
Human secreted pro  
protein involved i  
Arabidopsis thaliana  
Human polypeptide  
Zea mays protein f  
Drosophila melanog  
Arabidopsis thaliana  
Arabidopsis thaliana

PR 21-APR-1999;	99US-0130449.
PR 23-APR-1999;	99US-0130510.
PR 23-APR-1999;	99US-0130891.
PR 28-APR-1999;	99US-0131449.
PR 30-APR-1999;	99US-0132048.
PR 30-APR-1999;	99US-0132407.
PR 04-MAY-1999;	99US-0132484.
PR 05-MAY-1999;	99US-0132485.
PR 06-MAY-1999;	99US-0132486.
PR 07-MAY-1999;	99US-0132487.
PR 11-MAY-1999;	99US-0132863.
PR 14-MAY-1999;	99US-0134256.
PR 14-MAY-1999;	99US-0134218.
PR 14-MAY-1999;	99US-0134221.
PR 14-MAY-1999;	99US-0134370.
PR 18-MAY-1999;	99US-0134768.
PR 19-MAY-1999;	99US-0134941.
PR 20-MAY-1999;	99US-0135124.
PR 21-MAY-1999;	99US-0135353.
PR 24-MAY-1999;	99US-0135629.
PR 25-MAY-1999;	99US-0136021.
PR 27-MAY-1999;	99US-0136392.
PR 28-MAY-1999;	99US-0136787.
PR 01-JUN-1999;	99US-0137222.
PR 03-JUN-1999;	99US-0137528.
PR 04-JUN-1999;	99US-0137502.
PR 07-JUN-1999;	99US-0137724.
PR 08-JUN-1999;	99US-0138094.
PR 10-JUN-1999;	99US-0138540.
PR 10-JUN-1999;	99US-0138847.
PR 14-JUN-1999;	99US-0139119.
PR 16-JUN-1999;	99US-0139452.
PR 16-JUN-1999;	99US-0139453.
PR 17-JUN-1999;	99US-0139492.
PR 18-JUN-1999;	99US-0139454.
PR 18-JUN-1999;	99US-0139455.
PR 18-JUN-1999;	99US-0139456.
PR 18-JUN-1999;	99US-0139457.
PR 18-JUN-1999;	99US-0139458.
PR 18-JUN-1999;	99US-0139459.
PR 18-JUN-1999;	99US-0139460.
PR 18-JUN-1999;	99US-0139461.
PR 18-JUN-1999;	99US-0139462.
PR 18-JUN-1999;	99US-0139463.
PR 18-JUN-1999;	99US-0139750.
PR 18-JUN-1999;	99US-0139763.
PR 21-JUN-1999;	99US-0139817.
PR 22-JUN-1999;	99US-0139899.
PR 23-JUN-1999;	99US-0140353.
PR 23-JUN-1999;	99US-0140354.
PR 24-JUN-1999;	99US-0140695.
PR 28-JUN-1999;	99US-0140823.
PR 29-JUN-1999;	99US-0140991.
PR 30-JUN-1999;	99US-0141287.
PR 01-JUL-1999;	99US-0141842.
PR 01-JUL-1999;	99US-0142154.
PR 02-JUL-1999;	99US-0142055.
PR 06-JUL-1999;	99US-0142390.
PR 08-JUL-1999;	99US-0142803.
PR 09-JUL-1999;	99US-0142920.
PR 12-JUL-1999;	99US-0142977.
PR 13-JUL-1999;	99US-0143542.
PR 14-JUL-1999;	99US-0143624.
PR 15-JUL-1999;	99US-0144005.
PR 16-JUL-1999;	99US-0144085.
PR 16-JUL-1999;	99US-0144086.
PR 19-JUL-1999;	99US-0144325.
PR 19-JUL-1999;	99US-0144331.
PR 19-JUL-1999;	99US-0144332.
PR 19-JUL-1999;	99US-0144333.
PR 19-JUL-1999;	99US-0144334.
PR 19-JUL-1999;	99US-0144335.
PR 20-JUL-1999;	99US-0144352.
PR 20-JUL-1999;	99US-0144632.
PR 20-JUL-1999;	99US-0144684.
PR 21-JUL-1999;	99US-0144814.
PR 21-JUL-1999;	99US-0145086.
PR 21-JUL-1999;	99US-0145088.
PR 22-JUL-1999;	99US-0145085.
PR 22-JUL-1999;	99US-0145087.
PR 22-JUL-1999;	99US-0145089.
PR 22-JUL-1999;	99US-0145192.
PR 23-JUL-1999;	99US-0145145.
PR 23-JUL-1999;	99US-0145218.
PR 23-JUL-1999;	99US-0145224.
PR 26-JUL-1999;	99US-0145276.
PR 27-JUL-1999;	99US-0145913.
PR 27-JUL-1999;	99US-0145918.
PR 27-JUL-1999;	99US-0145919.
PR 28-JUL-1999;	99US-0145951.
PR 02-AUG-1999;	99US-0146386.
PR 02-AUG-1999;	99US-0146388.
PR 02-AUG-1999;	99US-0146389.
PR 03-AUG-1999;	99US-0147038.
PR 04-AUG-1999;	99US-0147204.
PR 04-AUG-1999;	99US-0147302.
PR 05-AUG-1999;	99US-0147192.
PR 05-AUG-1999;	99US-0147260.
PR 06-AUG-1999;	99US-0147303.
PR 06-AUG-1999;	99US-0147416.
PR 09-AUG-1999;	99US-0147493.
PR 09-AUG-1999;	99US-0147935.
PR 10-AUG-1999;	99US-0148171.
PR 11-AUG-1999;	99US-0148319.
PR 12-AUG-1999;	99US-0148341.
PR 13-AUG-1999;	99US-0148565.
PR 13-AUG-1999;	99US-0148684.
PR 16-AUG-1999;	99US-0149368.
PR 17-AUG-1999;	99US-0149175.
PR 18-AUG-1999;	99US-0149426.
PR 20-AUG-1999;	99US-0149722.
PR 20-AUG-1999;	99US-0149723.
PR 20-AUG-1999;	99US-0149929.
PR 23-AUG-1999;	99US-0149902.
PR 23-AUG-1999;	99US-0149930.
PR 25-AUG-1999;	99US-0150566.
PR 26-AUG-1999;	99US-0150884.
PR 27-AUG-1999;	99US-0151065.
PR 27-AUG-1999;	99US-0151066.
PR 27-AUG-1999;	99US-0151080.
PR 30-AUG-1999;	99US-0151303.
PR 31-AUG-1999;	99US-0151438.
PR 01-SEP-1999;	99US-0151930.
PR 07-SEP-1999;	99US-0152363.
PR 10-SEP-1999;	99US-0153070.
PR 13-SEP-1999;	99US-0153758.
PR 13-SEP-1999;	99US-0154018.
PR 16-SEP-1999;	99US-0154039.
PR 20-SEP-1999;	99US-0154779.
PR 22-SEP-1999;	99US-0155139.
PR 23-SEP-1999;	99US-0155486.
PR 24-SEP-1999;	99US-0155659.
PR 28-SEP-1999;	99US-0156458.
PR 29-SEP-1999;	99US-0156596.
PR 04-OCT-1999;	99US-0157117.
PR 05-OCT-1999;	99US-0157753.
PR 06-OCT-1999;	99US-0157865.
PR 07-OCT-1999;	99US-0158029.
PR 08-OCT-1999;	99US-0158232.
PR 12-OCT-1999;	99US-0158369.
PR 13-OCT-1999;	99US-0159293.
PR 13-OCT-1999;	99US-0159294.
PR 13-OCT-1999;	99US-0159295.
PR 14-OCT-1999;	99US-0159329.
PR 14-OCT-1999;	99US-0159330.

PR	14-OCT-1999	99US-015593.1
PR	14-OCT-1999	99US-015563.7
PR	14-OCT-1999	99US-015963.8
PR	18-OCT-1999	99US-015958.4
PR	21-OCT-1999	99US-016074.1
PR	21-OCT-1999	99US-016076.8
PR	21-OCT-1999	99US-016067.67
PR	21-OCT-1999	99US-016077.0
PR	21-OCT-1999	99US-016081.4
PR	21-OCT-1999	99US-016081.5
PR	22-OCT-1999	99US-016098.0
PR	22-OCT-1999	99US-016098.1
PR	22-OCT-1999	99US-016098.81
PR	25-OCT-1999	99US-016140.4
PR	25-OCT-1999	99US-016140.5
PR	25-OCT-1999	99US-016140.6
PR	25-OCT-1999	99US-016140.65
PR	26-OCT-1999	99US-016136.0
PR	26-OCT-1999	99US-016136.1
PR	28-OCT-1999	99US-016193.0
PR	28-OCT-1999	99US-016193.2
PR	28-OCT-1999	99US-016193.3
PR	29-OCT-1999	99US-016214.2

Query Match	73.1%	Score 19:	DB 21:	Length 378:
Best Local Similarity	40.0%	Pred. NO.	4.9e+02:	
Matches 4:	Conservative	0:	Mismatches 6:	Indels 0:
				Gaps 0

Qy	1	KKXXXXXXXXLD	10
Db	13	KKAADAAALD	22

RESULT 2  
AAB38172  
ID AAB38172 standard; Protein; 46 AA

AC	AAB38172;
XX	
DT	30-JAN-2001 (first entry)

DE Gene 3 human secreted protein homologous amino acid sequence #111.

KM Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;  
KM antihemetic; antiproliferative; cytostatic; cardiatic; vasotropic;  
KM cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;  
KM fungicide; ophthalmological; gene therapy; autoimmune disease; infection;  
KM hyperproliferative disorder; cardiovascular disorder; angiodemesis;  
KM cerebrovascular disorder; nervous system disorder; ocular disorder;  
KM wound healing; skin aging; food additive; preservative.

OS Homo sapiens.

PN WO200058468-A2

PD 05-OCT-2000.

PF 22-MAR-2000; 2000WO-US07526.

PR 26-MAR-1999; 99US-0126600.

PR 22-DEC-1999; 99US-0171550.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM, Komatsoulis G;

DR WPI; 2000-611713/58.

PT Nucleic acids encoding human secreted proteins, used to prevent, treat, ameliorate, or diagnose conditions such as autoimmune disorders, skin disorders and cancer -

XX The polynucleotide sequences given in AAC69399 to AAC69445 encode the human secreted proteins given in AAB38119 to AAB38165. AAB38166 to AAB38201 represent human secreted polypeptide sequences and proteins homologous to them, which are given in the exemplification of the present invention. Human secreted proteins have activities based on the tissues and cells the genes are expressed in. Example of activities include: immunosuppressive; antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant; vasotrophic; cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide; fungicide; and ophthalmologic. The polynucleotides and polypeptides can be are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Disorders which are diagnosed or treated include autoimmune diseases, hyperproliferative disorders, cardiovascular disorders, cerebrovascular disorders, anglogenesis, nervous system disorders, infections caused by bacteria, viruses and fungi and ocular disorders. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities. AAC69330 to AAC69398 and AAB38118 represent sequences used in the exemplification of the present invention.

**SQ Sequence 46 AA;**

Query Match	69.28	Score 18	DB 21	Length 46
Best Local Similarity	40.08	Pred No. 1.5e+02		
Matches 4	Conservative 0	Mismatches 6	Indels 0	Gaps 0

QY	1 KXXXXXXXLD 10
Db	19 KKSAAESMLD 28

RESULT 3  
AAB38173  
ID AAB38173 standard; Protein; 46 AA.

AC AAB38173;

DT 30-JAN-2001 (first entry)

DE Human secreted protein sequence encoded by gene 3 SEQ ID NO:112.

KM Human secreted protein; diagnosis; immunosuppressive; antitubercitic;  
 KM antineumatic; antiproliferative; cytostatic; cardant; vasotropic;  
 KM cerebroprotective; neurotropic; neuroprotective; antibacterial; virocidic;  
 KM fungicide; ophthalmological; gene therapy; autoimmune disease; infection;  
 KM hyperproliferative disorder; cardiovascular disorder; angiogenesis;  
 KM cerebrovascular disorder; nervous system disorder; ocular disorder;  
 KM wound healing; skin aging; food additive; preservative.

OS Homo sapiens.

PN WO200058468-A2

05-OCT-2000. PD

PF 22-MAR-2000; 2000WO-US07526.

PR 26-MAR-1999; 99US-0126600.

PR 22-DEC-1999; 99US-0171550.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM, Komatsoulis G;  
 ...

WPI; 2000-611713/58.

XX

PT Nucleic acids encoding human secreted proteins, used to prevent, treat, ameliorate, or diagnose conditions such as autoimmune disorders, skin disorders and cancer -

PS Disclosure: Page 362; 374pp; English.

XX The polynucleotide sequences given in AAC69399 to AAC69445 encode the  
CC human secreted proteins given in AAB38119 to AAB38165. AAB38166 to  
CC AAB38201 represent human secreted polypeptide sequences and proteins  
CC homologous to them, which are given in the exemplification of the present  
CC invention. Human secreted proteins have activities based on the tissues  
CC and cells the genes are expressed in. Example of activities include:  
CC immunosuppressive; antitumor; antirheumatic; antiproliferative;  
CC cytoprotective; cardiatic; vasotropic; cerebroprotective; neurotropic;  
CC neuroprotective; antibacterial; virucide; fungicide; and  
CC ophthalmological. The polynucleotides and polypeptides can be used to  
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
CC in diagnosing a pathological condition or susceptibility to a  
CC pathological condition. Disorders which are diagnosed or treated include  
CC autoimmune diseases, hyperproliferative disorders, cardiovascular  
CC disorders, cerebrovascular disorders, angioneurosis, nervous system  
CC disorders. The polypeptides can also be used to aid wound healing and  
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to  
CC maintain organs before transplantation, for supporting cell culture of  
CC primary tissues, to regenerate tissues and in chemotaxis. The  
CC polypeptides can also be used as a food additive or preservative and  
CC increase or decrease storage capabilities. AAC69399 to AAC69398 and  
CC AAB38118 represent sequences used in the exemplification of the present  
CC invention.

XX Sequence 46 AA;

Query Match 69.2%; Score 18; DB 21; Length 46;

Best Local Similarity 40.0%; Pred. No. 1.5e+02;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXXD 10  
||  
||  
Db 19 KKSAAESMLD 28

RESULT 4

AAG02207 ID AAG02207 standard; Protein: 147 AA.

XX AAG02207;

XX 06-OCT-2000 (first entry)

XX Human secreted protein, SEQ ID NO: 6288.

XX Human; 5' EST: expressed sequence tag; secreted protein; cDNA isolation;

KW gene therapy; chromosome mapping.

XX Homo sapiens.

XX EPI033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST ) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI: 2000-500381/45.

XX N-PSDB: AAC02213.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

PS Claim 13; SEQ ID 6288; 71pp + CD-ROM; English.

XX The present sequence is a polypeptide encoded by one of a large number  
CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
CC were prepared from total human RNAs or polyA+ RNAs derived from 30  
CC different tissues. EST sequences usually correspond mainly to the 3'  
CC untranslated region (UTR) of the mRNA because they are often obtained  
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
CC those cases where longer cDNA sequences have been obtained, the full 5'  
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
CC ends and can therefore be used to obtain full length cDNAs and genomic  
CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
CC chromosome mapping procedures. They are used to obtain upstream  
CC regulatory sequences and to design expression and secretion vectors.

XX Sequence 147 AA;

Query Match 69.2%; Score 18; DB 21; Length 147;

Best Local Similarity 40.0%; Pred. No. 4.4e+02;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXXD 10  
||  
||  
Db 78 KKKRAATTD 87

RESULT 5

AAM48271 ID AAM48271 standard; Protein: 152 AA.

XX AAM48271;

XX 23-JUN-1998 (first entry)

XX Rat. nujurin 1.

KW Rat; nujurin; cellular adhesion molecule; membrane bound; tumour;  
KW nerve injury induced gene; inflammation; nervous system.

XX Rattus sp.

XX WO9803650-A1.

XX 29-JAN-1998.

XX 24-JUL-1997; 97WO-US12210.

XX 24-JUL-1996; 96US-0672850.

XX (UNITV ) UNITV WASHINGTON.

XX Araki T, Mibrandt J;

XX WPI: 1998-120775/11.

XX N-PSDB: AAV20665.

XX New isolated nerve injury induced (nujurin) gene - used to develop  
XX products for treating conditions involving excessive or insufficient  
XX cellular adhesion, e.g. inflammation or tumours

XX Claim 5; Fig 1B; 80pp; English.

XX The present sequence represents rat nujurin 1 (nerve injury induced).  
XX The nujurin protein (NP) plays a role in axonal regeneration of  
XX peripheral nervous system (PNS) neuronal cells after injury. The  
XX products can be used for developing products for treating nujurin  
XX mediated disorders including conditions involving inappropriate (i.e.  
XX excessive or insufficient) cellular adhesion. Conditions involving

CC excessive cellular adhesion which may be treated include e.g.  
 CC inflammatory diseases such as rheumatoid arthritis, asthma, allergy  
 CC conditions, adult respiratory distress syndrome, inflammatory bowel  
 CC diseases (e.g. Crohn's disease, ulcerative colitis and regional  
 CC enteritis) and ophthalmic inflammatory diseases, autoimmune diseases,  
 CC thrombosis or inappropriate platelet aggregation conditions,  
 CC arteriosclerosis, reocclusion following thrombolysis, cardiovascular  
 CC metastasis conditions. The products can also be used to promote  
 CC cellular adhesion, e.g. in nerve regeneration, wound healing or  
 CC prosthetic implantation. The products can also be used for detection,  
 CC purification, diagnosis and screening assays.

CC Sequence 152 AA;

Query Match 69.2%; Score 18; DB 19; Length 152;

Best Local Similarity 40.0%; Pred. No. 4.5e+02; Mismatches 6; Indels 0; Gaps 0;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

1 KKKXXXXXLD 10  
 || ||  
 44 KKSAAESMD 53

DB

RESULT 6  
 AAM48272  
 ID AAM48272 standard; Protein: 152 AA.

XX AAM48272;

XX 23-JUN-1998 (first entry)

XX Human nlnjurin 1.

XX Human; nlnjurin; cellular adhesion molecule; membrane bound; tumour;  
 XX nerve injury induced gene; inflammation; nervous system.

XX Homo sapiens.

XX WO9803650-A1.

XX 29-JAN-1998.

XX 24-JUL-1997; 97WO-US12210.

XX 24-JUL-1996; 96US-0672850.

XX (UNITV ) UNITV WASHINGTON.

XX Araki T, Mlbrandt J;

XX WPI: 1998-120775/11.

XX N-PSDB; AAV20666.

XX New isolated nerve injury induced (nlnjurin) gene - used to develop  
 XX products for treating conditions involving excessive or insufficient  
 XX cellular adhesion, e.g. inflammation or tumours

XX Claim 6; Fig 2B; 80pp; English.

XX The present sequence represents human nlnjurin 1 (nerve injury induced).

XX The nlnjurin protein (NP) plays a role in axonal regeneration of  
 XX peripheral nervous system (PNS) neuronal cells after injury. The  
 XX products can be used for developing products for treating nlnjurin  
 XX mediated disorders including conditions involving inappropriate (i.e.  
 XX excessive or insufficient) cellular adhesion. Conditions involving  
 XX inflammatory diseases such as rheumatoid arthritis, asthma, allergy  
 XX conditions, adult respiratory distress syndrome, inflammatory bowel  
 XX diseases (e.g. Crohn's disease, ulcerative colitis and regional  
 XX enteritis) and ophthalmic inflammatory diseases, autoimmune diseases,  
 XX thrombosis or inappropriate platelet aggregation conditions,  
 XX arteriosclerosis, reocclusion following thrombolysis, cardiovascular

CC diseases, some forms of diabetes and neoplastic disease including  
 CC metastasis conditions. The products can also be used to promote  
 CC cellular adhesion, e.g. in nerve regeneration, wound healing or  
 CC prosthetic implantation. The products can also be used for detection,  
 CC purification, diagnosis and screening assays.

CC Sequence 152 AA;

Query Match 69.2%; Score 18; DB 19; Length 152;

Best Local Similarity 40.0%; Pred. No. 4.5e+02; Mismatches 6; Indels 0; Gaps 0;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

1 KKKXXXXXLD 10  
 || ||  
 44 KKSAAESMD 53

DB

RESULT 7  
 ABG20269  
 ID ABG20269 standard; Protein: 173 AA.

XX ABG20269;

XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #20260.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 XX food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

XX N-PSDB; AAS84456.

XX New isolated polynucleotide and encoded polypeptides, useful in  
 XX diagnostics, forensics, gene mapping, identification of mutations  
 XX responsible for genetic disorders or other traits and to assess  
 XX biodiversity

XX Claim 20; SEQ ID NO 50628; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
 XX polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 XX and gene mapping, and in recombinant production of (II). The  
 XX polynucleotides are also used in diagnostics as expressed sequence tags  
 XX for identifying expressed genes. (I) is useful in gene therapy techniques  
 XX to restore normal activity of (II) or to treat disease states involving  
 XX (II). (II) is useful for generating antibodies against it, detecting or  
 XX quantitating a polypeptide in tissue, as molecular weight markers and as  
 XX a food supplement. (I) and its binding partners are useful in medical  
 XX imaging of sites expressing (II). (I) and (II) are useful for treating  
 XX disorders involving aberrant protein expression or biological activity.

XX The polypeptide and polynucleotide sequences have applications in  
 XX diagnostics, forensics, gene mapping, identification of mutations  
 XX responsible for genetic disorders or other traits to assess biodiversity  
 XX and to produce other types of data and products dependent on DNA and  
 XX amino acid sequences. ABG00010-ABG30377 represent novel human  
 XX diagnostic amino acid sequences of the invention.



PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;  
 XX WPI; 1998-506364/43.  
 DR N-PSDB; AAV59686.  
 XX  
 PT New isolated human genes and the secreted polypeptide(s) they encode  
 PT - useful for diagnosis and treatment of e.g. cancers, neurological  
 PT disorders, immune diseases, inflammation or blood disorders  
 XX  
 PS Claim 1; Page 649-650, 721pp; English.  
 XX  
 CC This sequence represents a secreted human protein encoded by the nucleic  
 CC acid molecule designated Gene 176 from the human cDNA clone HPTBR48  
 CC (deposited as clone ATCC 97904 and ATCC 209050).  
 CC The gene can be used to generate fusion proteins by linking to the gene  
 CC to a human immunoglobulin Fc portion (e.g. AAV59502) for increasing the  
 CC stability of the fused protein as compared to the human protein only.  
 CC The invention relates to 186 novel genes and their fragments (nucleic  
 CC acid sequences: AAV59511-V59812; amino acid sequences AAW4731-W5026)  
 CC which are useful for preventing, treating or ameliorating medical  
 CC conditions e.g. by protein or gene therapy. Also, pathological  
 CC conditions can be diagnosed by determining the amount of the new  
 CC polypeptides in a sample or by determining the presence of mutations in  
 CC the new polynucleotides. Specific uses are described for each of the 186  
 CC polynucleotides, based on which tissues they are most highly expressed in  
 CC (see AAV59511 for described uses).  
 CC  
 SQ Sequence 200 AA;  
 XX  
 Query Match 69.2%; Score 18; DB 19; Length 200;  
 Best Local Similarity 40.0%; Pred. No. 5.8e+02;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 OY 1 KKKXXXXXLD 10  
 || ||  
 Db 151 KKAASHRAID 160  
 XX  
 RESULT 9  
 AAB47003  
 ID AAB47003 standard; Protein; 213 AA.  
 AC AAB47003;  
 XX  
 DT 22-MAR-2001 (first entry)  
 XX  
 DE Plant D-like cyclin inhibitor BRO2.  
 XX  
 KW Plant; D-like cyclin inhibitor gene; BRO4; hyperplastic; variant;  
 KW growth rate; dividing cells; inactivation; protoplast; seed;  
 KW root cell; meristem; leaf.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 FH Key Location/Qualifiers  
 FT Protein 20..147  
 FT /Label= BRO2  
 XX  
 PN WO200069883-A1.  
 XX  
 PD 23-NOV-2000.  
 XX  
 PF 15-MAY-2000; 2000WO-US13379.  
 XX  
 PR 14-MAY-1999; 99US-0134373.  
 XX  
 PA (HUTC-) HUTCHINSON CANCER RES CENT FRED.  
 XX  
 PI Roberts J, Kelly B;  
 XX  
 DR WPI; 2001-024998/03.  
 DR N-PSDB; AAC85203.  
 XX

PT Functionally inactivating expression of plant D-like cyclin inhibitor  
 PT gene for producing a hyperplastic variant plant, modulating the growth  
 PT and/or yield of plants, and increasing the proportion of dividing cells  
 XX  
 PS Example 1; Page 38-39; 50pp; English.  
 XX  
 CC The sequence given in AAB47003 represents a plant D1 cyclin  
 CC inhibitor, BRO2. The DNA encoding this sequence was isolated using  
 CC a yeast two hybrid screen. The BRO2 protein was found to contain a  
 CC seven amino acid sequence cyclin binding domain similar to that of  
 CC BRO1, BRO3 and BRO4 (see also AAB47005-6). The DNA encoding this  
 CC sequence is homologous to a sequence present in a D-like cyclin  
 CC inhibitor gene and when integrated at the corresponding locus,  
 CC functionally inactivates plant D-like cyclin inhibitor protein  
 CC expression. The BRO4 coding sequence may be used to produce a  
 CC hyperplastic variant plant, increase the growth rate of a plant,  
 CC or increase the proportion of dividing cells in a plant cell  
 CC population, relative to a wild-type plant, by functionally  
 CC inactivating the expression of a plant D-like cyclin inhibitor  
 CC gene in a plant. BRO4 is useful for increasing the proportion of  
 CC dividing cells in a plant cell population comprising protoplast,  
 CC seeds, root cells, meristem cells or leaf cells.  
 CC  
 SQ Sequence 213 AA;  
 XX  
 Query Match 69.2%; Score 18; DB 22; Length 213;  
 Best Local Similarity 40.0%; Pred. No. 6.1e+02;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 OY 1 KKKXXXXXLD 10  
 || ||  
 Db 10 KKTQYTSID 19  
 XX  
 RESULT 10  
 AAR29641  
 ID AAR29641 standard; Protein; 247 AA.  
 AC AAR29641;  
 XX  
 DT 05-FEB-1993 (first entry)  
 XX  
 DE PCTD ORF 6.  
 XX  
 KW CT; PCTD; epithelium; ocula mucosa; uro-genital mucosa; antigen;  
 KW monoclonal; polyclonal; antibody; vaccine.  
 XX  
 OS Chlamydia trachomatis.  
 XX  
 PN EP499681-A.  
 XX  
 PD 26-AUG-1992.  
 XX  
 PE 17-APR-1991; 91EP-0106110.  
 XX  
 PR 07-FEB-1991; 91IT-0000314.  
 XX  
 PA (ISTS) SCLAVO SPA.  
 XX  
 PI Comanducci M, Giuliani MM, Ratti G, Tecce MF;  
 XX  
 DR WPI; 1992-285922/35.  
 DR N-PSDB; AAQ27429.  
 XX  
 PT PCTD plasmid from Chlamydia Trachomatis and immunogenic proteins  
 PT - for diagnosing and vaccinating against Chlamydia infections  
 PT e.g. venereal lymphogranuloma  
 XX  
 PS Claim 1; Page 8-16; 40pp; English.  
 XX  
 CC The sequences given in AAR29636-43 are encoded by the plasmid isolated  
 CC from Chlamydia trachomatis (CT) serotype D, PCTD. This serotype  
 CC generally infects epithelial tissues, such as the ocular and

uro-genital mucous membranes, and shows a low virulence. Of the eight proteins encoded by the plasmid, seven are encoded by the sense strand and the eighth is encoded by the complementary strand. These proteins can be used as antigens for the preparation of poly- and mono-clonal antibodies to be used in diagnostics. The antigens can also be used in the formulation of vaccines against infections due to CT.

Sequence 247 AA:

Query Match 69.2%; Score 18; DB 13; Length 247;  
Best Local Similarity 40.0%; Pred. No. 7e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXXXXXXXXD 10  
11  
11  
DB 13 KKNQTAASID 22

RESULT 11

ABP41256 standard; Protein; 284 AA.

ABP41256;  
22-AUG-2002 (first entry)

Human ovarian antigen HOFNY16, SEQ ID NO:2388.

Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
ovarian cancer; breast cancer; tumour; reproductive system disorder;  
infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;  
inflammatory condition; immune disorder; blood disorder;  
cardiovascular disorder; respiratory disorder; neurological disorder;  
gastrointestinal disorder; urinary system disorder; drug screening;  
gene therapy; chromosome mapping; forensic analysis;  
antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
antiinflammatory; gynaecological; reproductive.

Homo sapiens.

WO200200677-A1.

03-JAN-2002.

07-JUN-2001; 2001WO-US18569.

07-JUN-2000; 2000US-209467P.

(HUMA-) HUMAN GENOME SCI INC.

Blrse CE, Rosen CA;

WPI: 2002-147878/19.

N-PSDB; AB054333.

Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
useful in the prevention, treatment and diagnosis of cancer (e.g.  
neurological diseases -

Claim 11; SEQ ID No 2388; 2922pp; English.

The invention relates to 2175 novel human ovarian antigens (ABP41054-  
ABP43228) and to cdnas encoding them (AB054131-AB056305), and also  
encompasses polypeptides 90% identical and polynucleotides 95% identical  
to the sequences of the invention. The invention additionally relates to  
recombinant vectors and host cells comprising human ovarian antigen  
polynucleotides, antibodies against human ovarian antigens, and the use  
of ovarian antigen polynucleotides and polypeptides in diagnosing,  
treating, prognosing or preventing various ovary and/or breast-related  
disorders. Such conditions include ovarian cancer and breast cancer, and  
metastatic tumours of ovarian or breast origin, reproductive system

disorders (e.g., infertility, disorders of pregnancy, anovulation,  
polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine  
disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
vaginitis), immune disorders (e.g., congenital and acquired  
immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
blood-related disorders (e.g., anaemia), cardiovascular disorders,  
respiratory disorders, neurological disorders, gastrointestinal disorders  
and urinary system disorders. Ovarian antigen polypeptides and  
polynucleotides may also be used in screening for compounds which  
modulate ovarian antigen expression or activity. The polynucleotides may  
further be used for gene therapy, chromosome mapping, in the  
identification of individuals and in forensic analysis, and the  
polypeptides may be used as food additives or to prepare antibodies  
useful in disease diagnosis, drug targeting and phenotyping. The present  
sequence represents a human ovarian antigen of the invention.  
Note: The sequence data for this patent did not form part of the printed  
specification, but was obtained in electronic format directly from WIPO  
at ftp.wipo.int/pub/published\_pcl\_sequences.

Sequence 284 AA:

Query Match 69.2%; Score 18; DB 23; Length 284;  
Best Local Similarity 40.0%; Pred. No. 7.9e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXXXXXXXXD 10  
11  
11  
DB 236 KKAASHRALD 245

RESULT 12

AAW93144 standard; Protein; 406 AA.

AAW93144;

21-MAY-1999 (first entry)

S. lividans xylanase A protein StmXlnA.

Xylanase A; acidophilic; thermostable; XYL I; XYL II; plant biomass;  
hemocellulase; beta-1,4 bond; xylosic chain; xylan; D-xylose; paper;  
pulp; chlorine bleaching; feed; beta-glucan; cellulose; lignin.

Streptomyces lividans.

US5871730-A.

16-FEB-1999.

29-JUL-1994; 94US-0282197.

29-JUL-1994; 94US-0282197.

(UTSH) UNITV SHERBROOKE.

Beaulieu C, Brzezinski R, Dery CV;

WPI: 1996-141348/15.

New acidophilic and thermostable xylanase enzymes from Actinomadura  
sp. FC7 - useful for treating plant biomass, especially paper and  
wood pulp, to degrade hemocellulose and hydrolyse xylan

Example 8; Fig 10; 60pp; English.

This invention describes the use of novel acidophilic and thermostable  
xylanase enzymes (XYL I and XYL II) from Actinomadura sp. FC7 which  
retain their activity under harsh industrial conditions (e.g. high  
temperature or wide pH ranges) and may be secreted by recombinant host  
cells, to treat plant biomass. Xylanases XYL I and XYL II are part of  
a large group of hemocellulase enzymes and function by cutting the



CC - beta-1,4 bonds within the xylosic chain of xylan (a polymer of D-xylose  
CC residues that is a major constituent of hemicellulose). This means that  
CC they may be used in the paper and pulp industry to improve the efficiency  
CC of the bleaching process by degrading the structure of the material.  
CC XYL I and XYL II may also be used to treat feed, by degrading a  
CC substrate with a high beta-glucan or cellulose content. XYL I and XYL II  
CC retain their activity at high temperatures (e.g. 70 deg. C) and at low  
CC pHs (e.g. 4.0), conditions which tend to denature most known xylanases.  
CC Enzymes that remain active in these conditions may be used in industrial  
CC processes that are carried out at high temperature and low pH to speed up  
CC other, non-enzymatic reactions, minimising costs, energy requirements,  
CC and the risk of pollution, (e.g. enzymes XYL I and XYL II can be used to  
CC facilitate chlorine bleaching of paper pulp which is carried out in hot,  
CC acidic conditions). Pretreatment with XYL I and XYL II, allows the  
CC bleaching agents to penetrate better, to remove lignin from the pulp and  
CC 'bleach' the colouration from it. This means smaller quantities of the  
CC agents can be used to produce the same or a better result. Also,  
CC disrupting the structure aids water drainage.  
CC NOTE: This patent is an equivalent to FI9503640.  
CC  
XX

SQ Sequence 406 AA;

Query Match 69.2%; Score 18; DB 17; Length 406;  
Best Local Similarity 40.0%; Pred. No. 1.1e+03;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXLD 10  
|| ||  
Db 330 KKAATYAVLD 339

RESULT 13  
AAW93155

ID AAW93155 standard; Protein; 420 AA.

XX AAW93155;

XX 21-MAY-1999 (first entry)

XX Streptomyces lividans xylanase A protein.

XX Xylanase A; acidophilic; thermostable; XYL I; XYL II; plant biomass;  
XX hemicellulase; beta-1,4 bond; xylosic chain; xylan; D-xylose; paper;  
XX pulp; chlorine bleaching; feed; beta-glucan; cellulose; lignin.

XX Streptomyces lividans.

XX US5871730-A.

XX 16-FEB-1999.

XX 29-JUL-1994; 94US-0282197.

XX 29-JUL-1994; 94US-0282197.

XX (UYSH ) UNIV SHERBROOKE.

XX Beaulieu C., Brzezinski R., Dery CV;

XX WPI; 1996-141348/15.

XX New acidophilic and thermostable xylanase enzymes from Actinomadura  
XX sp. FC7 - useful for treating plant biomass, especially paper and  
XX wood pulp, to degrade hemicellulose and hydrolyse xylan

XX Example 8; Fig 11; 60pp; English.

XX This invention describes the use of novel acidophilic and thermostable  
XX xylanase enzymes (XYL I and XYL II) from Actinomadura sp. FC7 which  
XX retain their activity under harsh industrial conditions (e.g. high  
XX temperature or wide pH ranges) and may be secreted by recombinant host  
XX cells, to treat plant biomass. Xylanases XYL I and XYL II are part of  
XX a large group of hemicellulase enzymes and function by cutting the

CC beta-1,4 bonds within the xylosic chain of xylan (a polymer of D-xylose  
CC residues that is a major constituent of hemicellulose). This means that  
CC they may be used in the paper and pulp industry to improve the efficiency  
CC of the bleaching process by degrading the structure of the material.  
CC XYL I and XYL II may also be used to treat feed, by degrading a  
CC substrate with a high beta-glucan or cellulose content. XYL I and XYL II  
CC retain their activity at high temperatures (e.g. 70 deg. C) and at low  
CC pHs (e.g. 4.0), conditions which tend to denature most known xylanases.  
CC Enzymes that remain active in these conditions may be used in industrial  
CC processes that are carried out at high temperature and low pH to speed up  
CC other, non-enzymatic reactions, minimising costs, energy requirements,  
CC and the risk of pollution, (e.g. enzymes XYL I and XYL II can be used to  
CC facilitate chlorine bleaching of paper pulp which is carried out in hot,  
CC acidic conditions). Pretreatment with XYL I and XYL II, allows the  
CC bleaching agents to penetrate better, to remove lignin from the pulp and  
CC 'bleach' the colouration from it. This means smaller quantities of the  
CC agents can be used to produce the same or a better result. Also,  
CC disrupting the structure aids water drainage.  
CC NOTE: This patent is an equivalent to FI9503640.  
CC  
XX

SQ Sequence 420 AA;

Query Match 69.2%; Score 18; DB 17; Length 420;  
Best Local Similarity 40.0%; Pred. No. 1.1e+03;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXLD 10  
|| ||  
Db 330 KKAATYAVLD 339

RESULT 14  
AAW93157

ID AAW93157 standard; Protein; 420 AA.

XX AAW93157;

XX 21-MAY-1999 (first entry)

XX Streptomyces lividans xylanase A protein.

XX Xylanase A; acidophilic; thermostable; XYL I; XYL II; plant biomass;  
XX hemicellulase; beta-1,4 bond; xylosic chain; xylan; D-xylose; paper;  
XX pulp; chlorine bleaching; feed; beta-glucan; cellulose; lignin.

XX Streptomyces lividans.

XX US5871730-A.

XX 16-FEB-1999.

XX 29-JUL-1994; 94US-0282197.

XX 29-JUL-1994; 94US-0282197.

XX (UYSH ) UNIV SHERBROOKE.

XX Beaulieu C., Brzezinski R., Dery CV;

XX WPI; 1996-141348/15.

XX New acidophilic and thermostable xylanase enzymes from Actinomadura  
XX sp. FC7 - useful for treating plant biomass, especially paper and  
XX wood pulp, to degrade hemicellulose and hydrolyse xylan

XX Example 9; Fig 13; 60pp; English.

XX This invention describes the use of novel acidophilic and thermostable  
XX xylanase enzymes (XYL I and XYL II) from Actinomadura sp. FC7 which  
XX retain their activity under harsh industrial conditions (e.g. high  
XX temperature or wide pH ranges) and may be secreted by recombinant host  
XX cells, to treat plant biomass. Xylanases XYL I and XYL II are part of  
XX a large group of hemicellulase enzymes and function by cutting the

CC beta-1,4 bonds within the xylosic chain of xylan (a polymer of D-xylose  
 CC residues that is a major constituent of hemicellulose). This means that  
 CC they may be used in the paper and pulp industry to improve the efficiency  
 CC of the bleaching process by degrading the structure of the material.  
 CC XYL I and XYL II may also be used to treat feed, by degrading a  
 CC substrate with a high beta-glucan or cellulose content. XYL I and XYL II  
 CC retain their activity at high temperatures (e.g. 70 deg. C) and at low  
 CC pHs (e.g. 4.0), conditions which tend to denature most known xylanases.  
 CC Enzymes that remain active in these conditions may be used in industrial  
 CC processes that are carried out at high temperature and low pH to speed up  
 CC other, non-enzymatic reactions, minimising costs, energy requirements,  
 CC and the risk of pollution, (e.g. enzymes XYL I and XYL II can be used to  
 CC facilitate chlorine bleaching of paper pulp which is carried out in hot,  
 CC acidic conditions). Pretreatment with XYL I and XYL II, allows the  
 CC bleaching agents to penetrate better, to remove lignin from the pulp and  
 CC 'bleach' the colouration from it. This means smaller quantities of the  
 CC agents can be used to produce the same or a better result. Also,  
 CC disrupting the structure aids water drainage.  
 CC NOTE: This patent is an equivalent to FI9503640.

SQ Sequence 420 AA;

Query Match 69.2%; Score 18; DB 17; Length 420;

Best Local Similarity 40.0%; Pred. No. 1.1e+03;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KXXXXXXD 10

DB 330 KKAAYTAVLD 339

#### RESULT 15

ABG19070 ABG19070 standard; Protein; 618 AA.

AC ABG19070;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #19061.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KV food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR MPI; 2001-639362/73.

DR N-PSDB; AAS83257.

PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity

PS Claim 20: SEQ ID NO 49429; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 618 AA;

Query Match 69.2%; Score 18; DB 22; Length 618;

Best Local Similarity 40.0%; Pred. No. 1.6e+03;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KXXXXXXD 10

DB 372 KKVTLSTSD 381

Search completed: June 5, 2003, 16:00:23  
 JOD time : 34.9394 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: June 5, 2003, 15:23:33 ; Search time 8.78788 Seconds  
(Without alignments)  
33.481 Million cell updates/sec

Title: US-09-150-947F-14  
Perfect score: 26  
Sequence: 1 KKKXXXXXLD 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*\n2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*\n3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*\n4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*\n5: /cgn2\_6/ptodata/1/1aa/6CUTS\_COMB.pep:\*\n6: /cgn2\_6/ptodata/1/1aa/backfillseq.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	69.2	100	4	US-08-672-850-13
2	18	69.2	152	4	US-08-672-850-4
3	18	69.2	152	4	US-08-672-850-6
4	18	69.2	152	4	US-08-672-850-12
5	18	69.2	200	4	US-09-149-476-495
6	18	69.2	237	2	US-08-933-750C-32
7	18	69.2	237	4	US-09-234-613-32
8	18	69.2	247	3	US-08-969-644-18
9	18	69.2	247	3	US-08-444-189-18
10	18	69.2	247	4	US-08-468-544-18
11	18	69.2	406	2	US-08-282-197C-52
12	18	69.2	420	2	US-08-282-197C-63
13	18	69.2	420	2	US-08-282-197C-66
14	18	69.2	491	2	US-08-468-812-8
15	18	69.2	491	4	US-08-590-563-8
16	18	69.2	705	2	US-08-663-566A-19
17	18	69.2	705	2	US-08-023-610-19
18	18	69.2	705	2	US-08-288-065A-19
19	18	69.2	705	2	US-08-362-240A-19
20	18	69.2	705	5	PCT-US95-10245-19
21	17	65.4	17	4	US-08-896-933-6
22	17	65.4	17	4	US-08-896-933-19
23	17	65.4	17	4	US-09-314-235-6
24	17	65.4	17	4	US-09-314-235-19
25	17	65.4	24	3	US-08-838-413A-20
26	17	65.4	24	3	US-08-838-413A-27
27	17	65.4	142	4	US-08-672-850-8

28	17	65.4	193	4	US-08-896-933-31	Sequence 31, App1
29	17	65.4	193	4	US-09-314-235-31	Sequence 31, App1
30	17	65.4	194	1	US-08-446-918A-6	Sequence 6, App1
31	17	65.4	194	1	US-08-580-806-6	Sequence 6, App1
32	17	65.4	224	1	US-08-152-456A-2	Sequence 2, App1
33	17	65.4	234	1	US-08-440-221-2	Sequence 2, App1
34	17	65.4	234	3	US-08-486-099-111	Sequence 11, App
35	17	65.4	234	3	US-08-360-107A-121	Sequence 121, App
36	17	65.4	234	3	US-08-484-223B-111	Sequence 111, App
37	17	65.4	234	3	US-08-919-597-111	Sequence 111, App
38	17	65.4	234	3	US-08-475-668A-111	Sequence 111, App
39	17	65.4	234	3	US-08-485-511A-111	Sequence 111, App
40	17	65.4	234	3	US-08-471-913A-111	Sequence 111, App
41	17	65.4	234	4	US-08-485-264A-111	Sequence 111, App
42	17	65.4	234	4	US-08-474-239A-111	Sequence 111, App
43	17	65.4	234	4	US-09-144-776B-12	Sequence 12, App1
44	17	65.4	238	4	US-08-896-933-28	Sequence 28, App1
45	17	65.4	238	4	US-09-314-235-28	Sequence 28, App1

## ALIGNMENTS

RESULT 1  
US-08-672-850-13  
Sequence 13, Application US/08672850  
Patent No. 6140117  
GENERAL INFORMATION:  
APPLICANT: Mubrandt, Jeffrey  
APPLICANT: Ataki, Toshiyuki  
TITLE OF INVENTION: NINURIN  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESS: Flehr, Hobach, Test, Albritton & Herbert  
STREET: Four Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: United States  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/672,850  
FILING DATE: 24-JUL-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Silva, Robin M.  
REGISTRATION NUMBER: 38,304  
REFERENCE/DOCKET NUMBER: A-63610  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 100 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-672-850-13

Query Match 69.2%; Score 18; DB 4; Length 100;  
Best Local Similarity 40.0%; Pred. No. 96;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXLD 10  
DB 44 KKAASMD 53

RESULT 2  
US-08-672-850-4  
Sequence 4, Application US/08672850  
Patent No. 6140117  
GENERAL INFORMATION:  
APPLICANT: Milbrandt, Jeffrey  
APPLICANT: ARAKI, Toshiyuki  
TITLE OF INVENTION: NINTURIN  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
STREET: Four Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: United States  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/672,850  
FILING DATE: 24-JUL-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Silva, Robin M.  
REGISTRATION NUMBER: 38,304  
REFERENCE/DOCKET NUMBER: A-63610  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-672-850-4

Query Match 69.2%; Score 18; DB 4; Length 152;  
Best Local Similarity 40.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKKXXXXXLD 10  
DB 44 KKSAAESMLD 53

RESULT 3  
US-08-672-850-6  
Sequence 6, Application US/08672850  
Patent No. 6140117  
GENERAL INFORMATION:  
APPLICANT: Milbrandt, Jeffrey  
APPLICANT: ARAKI, Toshiyuki  
TITLE OF INVENTION: NINTURIN  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
STREET: Four Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: United States  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/672,850

FILING DATE: 24-JUL-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Silva, Robin M.  
REGISTRATION NUMBER: 38,304  
REFERENCE/DOCKET NUMBER: A-63610  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-672-850-6

Query Match 69.2%; Score 18; DB 4; Length 152;  
Best Local Similarity 40.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKKXXXXXLD 10  
DB 44 KKSAAESMLD 53

RESULT 4  
US-08-672-850-12  
Sequence 12, Application US/08672850  
Patent No. 6140117  
GENERAL INFORMATION:  
APPLICANT: Milbrandt, Jeffrey  
APPLICANT: ARAKI, Toshiyuki  
TITLE OF INVENTION: NINTURIN  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
STREET: Four Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: United States  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/672,850  
FILING DATE: 24-JUL-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Silva, Robin M.  
REGISTRATION NUMBER: 38,304  
REFERENCE/DOCKET NUMBER: A-63610  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-672-850-12

Query Match 69.2%; Score 18; DB 4; Length 152;  
Best Local Similarity 40.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKKXXXXXLD 10

Fri Jun 6 09:14:32 2003

DB 44 KKSAAESMLD 53

RESULT 5

US-09-149-476-495

Sequence 495, Application US/09149476

Patent No. 6420526

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: 186 Human Secreted proteins

FILE REFERENCE: P2002P1

CURRENT APPLICATION NUMBER: US/09/149,476

CURRENT FILING DATE: 1998-09-08

EARLIER APPLICATION NUMBER: PCT/US98/04493

EARLIER FILING DATE: 1998-03-06

EARLIER APPLICATION NUMBER: 60/040,162

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,333

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/038,621

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,626

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,334

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,336

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,163

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/047,600

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,615

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,597

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,502

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,633

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,583

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,617

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,618

EARLIER FILING DATE: 1997-05-23

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/043,580

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,568

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,314

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,569

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,311

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,671

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,674

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,669

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,312

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,313

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,672

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,315

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/048,974

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/056,886

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,877

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,889

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,893

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,630

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,878

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,662

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,872

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,882

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,637

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,903

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,888

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,879

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,880

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,894

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,911

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,636

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,874

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,910

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,864

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,631

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,845

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,892

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/047,601

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/057,761  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/047,595  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,599  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,588  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,585  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,586  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,590  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,594  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,589  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,593  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,614  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,578  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,576  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/047,501  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,670  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/056,632  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,664  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,876  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,881  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,909  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,875  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,862  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,887  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,908  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/057,650  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/056,884  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057,669  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/049,610  
EARLIER FILING DATE: 1997-06-13  
EARLIER APPLICATION NUMBER: 60/061,060  
EARLIER FILING DATE: 1997-10-02

Query Match 69.2%; Score 18; DB 4; Length 200;  
Best Local Similarity 40.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXD 10  
DB 151 KKAASHRALD 160

RESULT 6  
US-08-933-750C-32  
Sequence 32, Application US/08933750C

Patent No. 5932442  
GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Bandman, Olga  
APPLICANT: Shah, Puri  
APPLICANT: Au-Young, Janice  
APPLICANT: Yue, Henry  
APPLICANT: Guegler, Karl J.  
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES  
NUMBER OF SEQUENCES: 98  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incycle Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/933,750C  
FILING DATE: September 23, 1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0356 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 237 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: BLANDNOT03  
CLONE: 1602473  
US-08-933-750C-32

Query Match 69.2%; Score 18; DB 2; Length 237;  
Best Local Similarity 40.0%; Pred. No. 2e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXD 10  
DB 189 KKAASHRALD 198

RESULT 7  
US-09-234-613-32  
Sequence 32, Application US/09234613  
Patent No. 6132973  
GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Bandman, Olga  
APPLICANT: Shah, Puri  
APPLICANT: Au-Young, Janice  
APPLICANT: Yue, Henry  
APPLICANT: Guegler, Karl J.  
APPLICANT: Corley, Neil C.  
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES

NUMBER OF SEQUENCES: 98  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/234,613  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/933,750  
FILING DATE: September 23, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0356 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 237 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: BLADNOT03  
CLONE: 1602473  
US-09-234-613-32

Query Match 69.2%; Score 18; DB 4; Length 237;  
Best Local Similarity 40.0%; Pred. No. 2e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKKXXXXXD 10  
DB 189 KKAASHRAD 198

RESULT 8  
US-08-969-644-18  
Sequence 18, Application US/08969644  
Patent No. 6096519  
GENERAL INFORMATION:  
APPLICANT: Rattl, Giulio  
APPLICANT: Comanducci, Maurizio  
APPLICANT: Tecce, Mario F.  
APPLICANT: Giuliani, Marzia M.  
TITLE OF INVENTION: PCID PLASMID ISOLATED FROM CHLAMYDIA  
TITLE OF INVENTION: TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY  
TITLE OF INVENTION: THEM: RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH  
STREET: 301 N. Washington Street  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22046-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/969,644  
FILING DATE: 13-NOV-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/467,152  
FILING DATE:  
APPLICATION NUMBER: US/07/661,820  
FILING DATE:  
APPLICATION NUMBER: IT MI 91A000314  
FILING DATE: 07-FEB-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Svensson, Leonard R.  
REGISTRATION NUMBER: 30,330  
REFERENCE/DOCKET NUMBER: 1267-202P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-241-1300  
TELEFAX: 703-241-2848  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 247 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-969-644-18

Query Match 69.2%; Score 18; DB 3; Length 247;  
Best Local Similarity 40.0%; Pred. No. 2e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKKXXXXXD 10  
DB 13 KKNOTRASLD 22

RESULT 9  
US-08-444-189-18  
Sequence 18, Application US/08444189  
Patent No. 6110705  
GENERAL INFORMATION:  
APPLICANT: Rattl, Giulio  
APPLICANT: Comanducci, Maurizio  
APPLICANT: Tecce, Mario F.  
APPLICANT: Giuliani, Marzia M.  
TITLE OF INVENTION: PCID PLASMID ISOLATED FROM CHLAMYDIA  
TITLE OF INVENTION: TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY  
TITLE OF INVENTION: THEM: RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH  
STREET: 301 N. Washington Street  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22046-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/444,189  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/180,528  
FILING DATE:  
APPLICATION NUMBER: US/07/991,512  
FILING DATE:  
APPLICATION NUMBER: US/07/661,820

FILING DATE:  
 APPLICATION NUMBER: IT MI 91A000314  
 FILING DATE: 07-FEB-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Svensson, Leonard R.  
 REGISTRATION NUMBER: 30,330  
 REFERENCE/DOCKET NUMBER: 1267-202P  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 703-241-1300  
 TELEFAX: 703-241-2848  
 TELETYPE: 248345  
 INFORMATION FOR SEQ ID NO: 18:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 247 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-444-189-18

Query Match 69.2%; Score 18; DB 3; Length 247;  
 Best Local Similarity 40.0%; Pred. No. 2e+02;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXLD 10  
 || ||  
 DB 13 KKNQTAASLD 22

RESULT 10  
 US-08-468-544-18  
 Sequence 18, Application US/08468544  
 Patent No. 6248563  
 GENERAL INFORMATION:  
 APPLICANT: Ratti, Giulio  
 APPLICANT: Comanducci, Maurizio  
 APPLICANT: Tecce, Mario F.  
 APPLICANT: Giuliani, Marzia M.  
 TITLE OF INVENTION: PCTD PLASMID ISOLATED FROM CHILAMYDIA  
 TITLE OF INVENTION: TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY  
 TITLE OF INVENTION: THEM; RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID  
 NUMBER OF SEQUENCES: 23  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH  
 STREET: 301 N. Washington Street  
 CITY: Falls Church  
 STATE: Virginia  
 COUNTRY: USA  
 ZIP: 22046-0747  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/468,544  
 FILING DATE: 06-JUN-1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/661,820  
 FILING DATE: 28-FEB-1991  
 APPLICATION NUMBER: IT MI 91A000314  
 FILING DATE: 07-FEB-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Svensson, Leonard R.  
 REGISTRATION NUMBER: 30,330  
 REFERENCE/DOCKET NUMBER: 1267-202P  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 703-241-1300  
 TELEFAX: 703-241-2848  
 TELETYPE: 248345  
 INFORMATION FOR SEQ ID NO: 18:  
 SEQUENCE CHARACTERISTICS:

LENGTH: 247 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-468-544-18

Query Match 69.2%; Score 18; DB 4; Length 247;  
 Best Local Similarity 40.0%; Pred. No. 2e+02;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXLD 10  
 || ||  
 DB 13 KKNQTAASLD 22

RESULT 11  
 US-08-282-197C-52  
 Sequence 52, Application US/08282197C  
 Patent No. 5871730  
 GENERAL INFORMATION:  
 APPLICANT: Brzezinski, Ryszard  
 APPLICANT: Dery, Claude V  
 APPLICANT: Beaulieu, Carole  
 TITLE OF INVENTION: Thermostable Xylanase DNA, Protein and  
 TITLE OF INVENTION: Methods of Use  
 NUMBER OF SEQUENCES: 67  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
 STREET: 1100 New York Ave., NW  
 CITY: Washington  
 STATE: DC  
 COUNTRY: USA  
 ZIP: 20005  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/282,197C  
 FILING DATE: 29-JUL-1994  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Cimbal, Michele A  
 REGISTRATION NUMBER: 33,851  
 REFERENCE/DOCKET NUMBER: 1050.0410000  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-371-2600  
 TELEFAX: 202-371-2540  
 INFORMATION FOR SEQ ID NO: 52:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 406 amino acids  
 TYPE: amino acid  
 TOPOLOGY: both  
 US-08-282-197C-52

Query Match 69.2%; Score 18; DB 2; Length 406;  
 Best Local Similarity 40.0%; Pred. No. 3.1e+02;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXLD 10  
 || ||  
 DB 330 KKAAYTAVLD 339

RESULT 12  
 US-08-282-197C-63  
 Sequence 63, Application US/08282197C  
 Patent No. 5871730  
 GENERAL INFORMATION:  
 APPLICANT: Brzezinski, Ryszard  
 APPLICANT: Dery, Claude V  
 APPLICANT: Beaulieu, Carole



TITLE OF INVENTION: Thermostable Xylanase DNA, Protein and  
TITLE OF INVENTION: Methods of Use  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
STREET: 1100 New York Ave., NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/282,197C  
FILING DATE: 29-JUL-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Cimbalala, Michele A  
REGISTRATION NUMBER: 33,851  
REFERENCE/DOCKET NUMBER: 1050.0410000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 63:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 420 amino acids  
TYPE: amino acid  
TOPOLOGY: both  
US-08-282-197C-63

Query Match 69.2%; Score 18; DB 2; Length 420;  
Best Local Similarity 40.0%; Pred. No. 3.1e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKKXXXXLD 10  
DB 330 KKAATYAVLD 339

RESULT 13  
US-08-282-197C-66  
Sequence 66, Application US/08282197C  
Patent No. 5871730  
GENERAL INFORMATION:  
APPLICANT: Brzezinski, Ryszard  
APPLICANT: Dery, Claude V  
APPLICANT: Beaulieu, Carole  
TITLE OF INVENTION: Thermostable Xylanase DNA, Protein and  
TITLE OF INVENTION: Methods of Use  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
STREET: 1100 New York Ave., NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/282,197C  
FILING DATE: 29-JUL-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Cimbalala, Michele A  
REGISTRATION NUMBER: 33,851  
REFERENCE/DOCKET NUMBER: 1050.0410000

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 66:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 420 amino acids  
TYPE: amino acid  
TOPOLOGY: both  
US-08-282-197C-66

Query Match 69.2%; Score 18; DB 2; Length 420;  
Best Local Similarity 40.0%; Pred. No. 3.1e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKKXXXXLD 10  
DB 330 KKAATYAVLD 339

RESULT 14  
US-08-468-812-8  
Sequence 8, Application US/08468812  
Patent No. 5935836  
GENERAL INFORMATION:  
APPLICANT: Vehmaenper, Jari  
APPLICANT: M ntyl, Arja  
APPLICANT: Fagerstr m, Richard  
APPLICANT: Lantto, Raija  
APPLICANT: Paloheimo, Marja  
APPLICANT: Suominen, Pirkko  
APPLICANT: Lahtinen, Tarja  
APPLICANT: Kristo, Paula  
TITLE OF INVENTION: Actinomadura Xylanase Sequences and Methods  
TITLE OF INVENTION: of Use  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX  
STREET: 1100 New York Ave., N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,812  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/332,412  
FILING DATE: 31-OCT-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/282,001  
FILING DATE: 29-JUL-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Bugalsky, Larry B.  
REGISTRATION NUMBER: 35,086  
REFERENCE/DOCKET NUMBER: 1050.0340002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 491 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

POSITION IN GENOME:  
CHROMOSOME/SEGMENT: M64551  
US-08-468-812-8

Query Match  
Best Local Similarity 40.0%; Score 18; DB 2; Length 491;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKKXXXXXLD 10  
DB 333 KKAAYTAVLD 342

Query Match  
Best Local Similarity 40.0%; Score 18; DB 4; Length 491;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKKXXXXXLD 10  
DB 333 KKAAYTAVLD 342

Search completed: June 5, 2003, 15:32:43  
Job time: 9.78788 secs

RESULT 15  
US-08-590-563-8

Sequence 8, Application US/08590563  
Patent No. 6300114

## GENERAL INFORMATION:

APPLICANT: M ntyl, Arja  
APPLICANT: Veinmaa, Jari  
APPLICANT: Fagerstr m, Richard  
APPLICANT: Lantto, Raija  
APPLICANT: Paloheimo, Marja  
APPLICANT: Suominen, Pirkko  
APPLICANT: Lahtinen, Tarja  
TITLE OF INVENTION: Production and Secretion of Proteins of  
NUMBER OF SEQUENCES: 39  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C.  
STREET: 1100 New York Ave., N.W. Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/590,563  
FILING DATE: 26-JAN-1996  
CLASSIFICATION: 536

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/468,812  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/332,412  
FILING DATE: 31-OCT-1994  
CLASSIFICATION: 536

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/282,001  
FILING DATE: 29-JUL-1994  
CLASSIFICATION: 536

## ATTORNEY/AGENT INFORMATION:

NAME: Bugalsky, Lawrence B.  
REGISTRATION NUMBER: 35,086  
REFERENCE/DOCKET NUMBER: 1050.0340003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:

LENGTH: 491 amino acids  
TYPE: amino acid

STRANDEDNESS: not relevant  
TOPOLOGY: linear

MOLECULE TYPE: peptide  
POSITION IN GENOME:

CHROMOSOME/SEGMENT: M64551  
US-08-590-563-8

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:31:01 ; Search time 12.1212 Seconds  
(without alignments)  
85.173 Million cell updates/sec

Title: US-09-150-947F-14

Sequence: 1 KRXXXXXXXXD 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published\_Applications\_AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PTCT\_NEW\_PUB pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PTCTUS\_PUBCOMB pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	18	69.2	200	9	US-09-809-391-495
2	18	69.2	237	10	US-09-840-787-32
3	18	69.2	365	9	US-10-125-692-16
4	18	69.2	491	10	US-09-770-621-8
5	17	65.4	12	9	US-10-113-809-3
6	17	65.4	58	12	US-10-001-879-145
7	17	65.4	107	9	US-09-465-714-4
8	17	65.4	142	10	US-09-729-674-76
9	17	65.4	151	9	US-09-465-714-3
10	17	65.4	194	1	US-08-882-431-12
11	17	65.4	233	1	US-08-882-431-12
12	17	65.4	234	9	US-09-870-759-18
13	17	65.4	234	9	US-09-870-759-18
14	17	65.4	238	9	US-10-002-784A-12
15	17	65.4	266	1	US-08-882-431-14
16	17	65.4	266	9	US-09-870-759-12
17	17	65.4	266	9	US-10-002-784A-14
18	17	65.4	383	9	US-09-738-626-6821
19	17	65.4	490	10	US-09-841-132-434

20	17	65.4	582	9	US-10-046-938-29	Sequence 29, Appl
21	17	65.4	623	9	US-09-738-626-5990	Sequence 5990, Ap
22	16	61.5	10	10	US-09-150-947B-3	Sequence 3, Appl1
23	16	61.5	10	10	US-09-150-947B-4	Sequence 4, Appl1
24	16	61.5	12	9	US-10-113-809-1	Sequence 1, Appl1
25	16	61.5	12	9	US-10-113-809-4	Sequence 4, Appl1
26	16	61.5	12	10	US-09-150-947B-2	Sequence 2, Appl1
27	16	61.5	13	10	US-09-150-947B-5	Sequence 5, Appl1
28	16	61.5	13	10	US-09-150-947B-6	Sequence 6, Appl1
29	16	61.5	13	10	US-09-150-947B-11	Sequence 11, Appl
30	16	61.5	14	10	US-09-150-947B-9	Sequence 9, Appl1
31	16	61.5	14	10	US-09-150-947B-10	Sequence 10, Appl1
32	16	61.5	16	10	US-09-150-947B-7	Sequence 7, Appl1
33	16	61.5	25	10	US-09-803-165-29	Sequence 29, Appl
34	16	61.5	26	10	US-09-150-947B-8	Sequence 8, Appl1
35	16	61.5	35	10	US-09-150-947B-8	Sequence 8, Appl1
36	16	61.5	36	9	US-09-764-891-4313	Sequence 4313, Ap
37	16	61.5	38	10	US-09-071-838-150	Sequence 150, App
38	16	61.5	39	10	US-09-843-845-12	Sequence 12, Appl
39	16	61.5	63	9	US-10-044-692-65	Sequence 65, Appl
40	16	61.5	63	9	US-10-044-539-65	Sequence 65, Appl
41	16	61.5	69	9	US-09-843-676-16	Sequence 16, Appl
42	16	61.5	69	9	US-09-766-253-16	Sequence 16, Appl
43	16	61.5	69	9	US-09-438-486-16	Sequence 16, Appl
44	16	61.5	69	9	US-10-053-758-16	Sequence 16, Appl
45	16	61.5	69	9	US-10-054-295-16	Sequence 16, Appl

#### ALIGNMENTS

RESULT 1  
US-09-809-391-495  
; Sequence 495, Application US/09809391  
; Publication No. US20030049618A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 186 Human Secreted proteins  
; FILE REFERENCE: P2002P2  
; CURRENT APPLICATION NUMBER: US/09/809,391  
; CURRENT FILING DATE: 2001-03-16  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 761  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 495  
; LENGTH: 200  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (200)  
; OTHER INFORMATION: Xaa equals stop translation  
US-09-809-391-495

Query Match 69.2%; Score 18; DB 9; Length 200;  
Best Local Similarity 40.0%; Pred. No. 2e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KRXXXXXXXXD 10  
DB 151 KRAASHRAID 160

RESULT 2  
US-09-840-787-32  
; Sequence 32, Application US/09840787  
; Patent No. US20020058264A1  
; GENERAL INFORMATION:  
; APPLICANT: Lal, Preeti  
; Hillman, Jennifer L.  
; Bandman, Olga  
; Shah, Purvi  
; Au-Young, Janice

```

Yue, Henry
Guegler, Karl J.
Corley, Neil C.
TITLE OF INVENTION: HDMAN REGULATORY MOLECULES
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/840,787
FILING DATE: 23-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/518,865
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 237 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BLADNOT03
CLONE: 1602473
US-09-840-787-32
SEQUENCE DESCRIPTION: SEQ ID NO: 32 :
Query Match 69.2% Score 18; DB 10; Length 237;
Best Local Similarity 40.0% Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0
QY 1 KKKKXXXXLD 10
||
||
Db 189 KKAASHRALD 198

```

```

: LENGTH: 365
: TYPE: PRT
: ORGANISM: P. mirabilis
US-10-125-692-16
Query Match 69.2%; Score 18; DB 9; Length 365;
Best Local Similarity 40.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0.
OY 1 KKKXXXXXLD 10
Db 197 KKIINAATLD 206

RESULT 4
US-09-770-621-8
: Sequence 8, Application US/09770621
: Patent No. US20010024815A1
: GENERAL INFORMATION:
APPLICANT: M nyl, Arja
APPLICANT: Vehmaaper, Jari
APPLICANT: Fagerstr m, Richard
APPLICANT: Lanto, Raija
APPLICANT: Paloheimo, Marja
APPLICANT: Suominen, Pirkko
APPLICANT: Iahinen, Tarja
TITLE OF INVENTION: Production and Secretion of Proteins of
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESS: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 New York Ave., N.W. Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/770,621
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/590,563
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/332,412
FILING DATE: 31-OCT-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/282,001
FILING DATE: 29-JUL-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Bugalsky, Lawrence B.
REGISTRATION NUMBER: 35,086
REFERENCE/DOCKET NUMBER: 1050, 0340003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 491 amino acids
TYPE: amino acid
STRANDEDNESS: NO. US20010024815A1 Relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
POSITION IN GENOME:
CHROMOSOME/SEGMENT: M64551
US-09-770-621-8

```

Query Match 69.2%; Score 18; DB 10; Length 491;  
Best Local Similarity 40.0%; Pred. No. 4.5e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXXXXXXD 10  
DB 333 KKAAYTAVLD 342

RESULT 5  
US-10-113-809-3  
; Sequence 3, Application US/10113809  
; Patent No. US2002017573A1  
; GENERAL INFORMATION:  
; APPLICANT: Maschke, Hans, E.  
; TITLE OF INVENTION: EXOTOXIN-LIGAND  
; FILE REFERENCE: MBP-011XX  
; CURRENT APPLICATION NUMBER: US/10/113,809  
; CURRENT FILING DATE: 2002-03-29  
; PRIOR APPLICATION NUMBER: 101 16 042.9-41  
; PRIOR FILING DATE: 2001-03-30  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Bacterial Toxin Binding Ligand  
US-10-113-809-3

Query Match 65.4%; Score 17; DB 9; Length 12;  
Best Local Similarity 40.0%; Pred. No. 33;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXXXXXXD 10  
DB 3 KQOLAISTLD 12

RESULT 6  
US-10-001-879-145  
; Sequence 145, Application US/10001879  
; Patent No. US20020127237A1  
; GENERAL INFORMATION:  
; APPLICANT: Salceda, Susana  
; APPLICANT: Macina, Roberto  
; APPLICANT: Recipon, Hervé  
; APPLICANT: Caferkey, Robert  
; APPLICANT: Ali, Shujath  
; APPLICANT: Sun, Yongming  
; APPLICANT: Liu, Chenghua  
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and  
; FILE REFERENCE: DEX-0281  
; CURRENT APPLICATION NUMBER: US/10/001,879  
; CURRENT FILING DATE: 2001-11-20  
; PRIOR APPLICATION NUMBER: 60/252,188  
; PRIOR FILING DATE: 2000-11-21  
; NUMBER OF SEQ ID NOS: 201  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 145  
; LENGTH: 58  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-001-879-145

Query Match 65.4%; Score 17; DB 12; Length 58;  
Best Local Similarity 40.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 1 KXXXXXXD 10

DB 42 KKRTOATLD 51

RESULT 7  
US-09-465-714-4  
; Sequence 4, Application US/09465714  
; Publication No. US20030032582A1  
; GENERAL INFORMATION:  
; APPLICANT: Wahlsten, Jennifer L.  
; APPLICANT: Ramakrishnan, S.  
; APPLICANT: Schlievert, Patrick M.  
; TITLE OF INVENTION: SPONTANEOUS MEMBRANE INSERTION  
; FILE REFERENCE: 09531/003001  
; CURRENT APPLICATION NUMBER: US/09/465,714  
; CURRENT FILING DATE: 1999-12-17  
; PRIOR APPLICATION NUMBER: 09/001,593  
; PRIOR FILING DATE: 1998-12-31  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 107  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Truncated bacterial sequence  
US-09-465-714-4

Query Match 65.4%; Score 17; DB 9; Length 107;  
Best Local Similarity 40.0%; Pred. No. 2.4e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXXXXXXD 10  
DB 34 KQOLAISTLD 43

RESULT 8  
US-09-729-674-76  
; Sequence 76, Application US/09729674  
; Patent No. US20010039335A1  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Kenneth  
; APPLICANT: McCoy, John M.  
; APPLICANT: LaVallie, Edward R.  
; APPLICANT: Collins-Racie, Lisa A.  
; APPLICANT: Evans, Cheryl  
; APPLICANT: Werberg, David  
; APPLICANT: Treacy, Maurice  
; APPLICANT: Agostino, Michael J.  
; APPLICANT: Steindinger II, Robert J.  
; APPLICANT: Spaulding, Vikki  
; APPLICANT: Wong, Gordon G.  
; APPLICANT: Clark, Hilary  
; APPLICANT: Fetschel, Kim  
; APPLICANT: Genetics Institute, Inc.  
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYPEPTIDES ENCODING THEM  
; FILE REFERENCE: 6055-64X  
; CURRENT APPLICATION NUMBER: US/09/729,674  
; CURRENT FILING DATE: 2000-12-04  
; PRIOR APPLICATION NUMBER: 09/539,330  
; PRIOR FILING DATE: 2000-03-30  
; NUMBER OF SEQ ID NOS: 283  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 76  
; LENGTH: 142  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-729-674-76

Query Match 65.4%; Score 17; DB 10; Length 142;  
Best Local Similarity 40.0%; Pred. No. 3.1e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXKD 10  
DB 30 KKSVAESMID 39

US-09-465-714-2

Sequence 2, Application US/09465714  
Publication No. US20030032582A1  
GENERAL INFORMATION:  
APPLICANT: Wahlsten, Jennifer L.  
APPLICANT: Ramakrishnan, S.  
APPLICANT: Schlievert, Patrick M.  
TITLE OF INVENTION: SPONTANEOUS MEMBRANE INSERTION  
FILE REFERENCE: 09531/003001  
CURRENT APPLICATION NUMBER: US/09/465,714  
CURRENT FILING DATE: 1999-12-17  
PRIOR APPLICATION NUMBER: 09/001,593  
PRIOR FILING DATE: 1998-12-31  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 151  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Hybrid containing bacterial and mammalian sequence  
US-09-465-714-2

Query Match  
Best Local Similarity 40.0%; Score 17; DB 9; Length 151;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXKD 10  
DB 34 KQOLAISTLD 43

US-09-465-714-3  
Sequence 3, Application US/09465714  
Publication No. US20030032582A1  
GENERAL INFORMATION:  
APPLICANT: Wahlsten, Jennifer L.  
APPLICANT: Ramakrishnan, S.  
APPLICANT: Schlievert, Patrick M.  
TITLE OF INVENTION: SPONTANEOUS MEMBRANE INSERTION  
FILE REFERENCE: 09531/003001  
CURRENT APPLICATION NUMBER: US/09/465,714  
CURRENT FILING DATE: 1999-12-17  
PRIOR APPLICATION NUMBER: 09/001,593  
PRIOR FILING DATE: 1998-12-31  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 194  
TYPE: PRT  
ORGANISM: Bacterial Sequence  
US-09-465-714-3

Query Match  
Best Local Similarity 40.0%; Score 17; DB 9; Length 194;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXKD 10  
DB 121 KQOLAISTLD 130

US-08-882-431-12  
Sequence 12, Application US/08882431  
Publication No. US20030009015A1

GENERAL INFORMATION:  
APPLICANT: Robert G. Ulrich,  
APPLICANT: Mark A. Olson  
APPLICANT: Sina Bavari  
TITLE OF INVENTION: Bacterial Superantigen  
TITLE OF INVENTION: Vaccines  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John Moran  
STREET: US Army MRC -504 Scott Street MCMR-JA (John Moran-Patent Atty)  
CITY: FORT DETRICK  
STATE: MARYLAND  
COUNTRY: USA  
ZIP: 21702-5012

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: Macintosh 7.5  
SOFTWARE: Microsoft Word 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/882,431  
FILING DATE: June 25, 1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Moran, John  
REGISTRATION NUMBER: 26,313  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 619-2065  
TELEFAX: (301) 619-7714  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 233  
TYPE: Amino Acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
US-08-882-431-12

Query Match  
Best Local Similarity 40.0%; Score 17; DB 1; Length 233;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXKD 10  
DB 160 KQOLAISTLD 169

US-09-870-759-18  
Sequence 18, Application US/09870759  
Patent No. US20020177551A1  
GENERAL INFORMATION:  
APPLICANT: TERMAN, David S  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE  
FILE REFERENCE: 870759  
CURRENT APPLICATION NUMBER: US/09/870,759  
CURRENT FILING DATE: 2002-01-14  
PRIOR APPLICATION NUMBER: US 60/208,128  
PRIOR FILING DATE: 2000-05-30  
NUMBER OF SEQ ID NOS: 166  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 18  
LENGTH: 234  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-870-759-18

Query Match  
Best Local Similarity 40.0%; Score 17; DB 9; Length 234;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXXXXXXLD 10  
||  
Db 161 KKOLAISTLD 170

RESULT 13  
US-10-002-784A-12  
; Sequence 12, Application US/10002784A  
; Publication No. US20030036644A1  
; GENERAL INFORMATION:  
; /33  
; APPLICANT: Ulrich, Robert G.  
; TITLE OF INVENTION: Bacterial Superantigen Vaccines  
; FILE REFERENCE: 003/233/SAP  
; CURRENT APPLICATION NUMBER: US/10/002,784A  
; CURRENT FILING DATE: 2001-11-26  
; PRIOR APPLICATION NUMBER: 08/882,431; 09/144,776  
; PRIOR FILING DATE: 97-06-25; 98-09-01  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: Apple Macintosh Microsoft Word 6.0  
; SEQ ID NO 12  
; LENGTH: 234  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: toxin shock syndrom toxin-1 mutant  
US-10-002-784A-12

Query Match 65.4%; Score 17; DB 9; Length 234;  
Best Local Similarity 40.0%; Pred. No. 4.9e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXXXXXXLD 10  
||  
Db 161 KKOLAISTLD 170

RESULT 14  
US-09-465-714-1  
; Sequence 1, Application US/09465714  
; Publication No. US20030032582A1  
; GENERAL INFORMATION:  
; APPLICANT: Wahlsten, Jennifer L.  
; APPLICANT: Ramakrishnan, S.  
; TITLE OF INVENTION: SPONTANEOUS MEMBRANE INSERTION  
; FILE REFERENCE: 09531/003001  
; CURRENT APPLICATION NUMBER: US/09/465,714  
; CURRENT FILING DATE: 1999-12-17  
; PRIOR APPLICATION NUMBER: 09/001,593  
; PRIOR FILING DATE: 1998-12-31  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 238  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Hybrid containing bacterial and mammalian sequence  
US-09-465-714-1

Query Match 65.4%; Score 17; DB 9; Length 238;  
Best Local Similarity 40.0%; Pred. No. 5e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXXXXXXLD 10  
||  
Db 121 KKOLAISTLD 130

RESULT 15

US-08-882-431-14  
; Sequence 14, Application US/0882431  
; Publication No. US20030009015A1  
; GENERAL INFORMATION:  
; APPLICANT: Robert G. Ulrich,  
; APPLICANT: Mark A. Olson  
; APPLICANT: Sina Bavari  
; TITLE OF INVENTION: Bacterial Superantigen  
; TITLE OF INVENTION: Vaccines  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: John Moran  
; STREET: US Army MRC -504 Scott Street MCMR-JA (John Moran-Patent Atty)  
; CITY: FORT DETRICK  
; STATE: MARYLAND  
; COUNTRY: USA  
; ZIP: 21702-5012  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Macintosh 7.5  
; SOFTWARE: Microsoft Word 6.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/882,431  
; FILING DATE: June 25, 1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Moran, John  
; REGISTRATION NUMBER: 26,313  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 619-2065  
; TELEFAX: (301) 619-7714  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 266  
; TYPE: Amino Acid  
; STRANDEDNESS: Unknown  
; TOPOLOGY: Unknown  
; MOLECULE TYPE: Peptide  
US-08-882-431-14

Query Match 65.4%; Score 17; DB 1; Length 266;  
Best Local Similarity 40.0%; Pred. No. 5.5e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXXXXXXLD 10  
||  
Db 178 KSVTAQELD 187

Search completed: June 5, 2003, 15:49:03  
Job time: 13.1212 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:24:01 ; Search time 9.84848 seconds  
(without alignments)  
97.613 Million cell updates/sec

Title: US-09-150-947F-14

Perfect score: 26

Sequence: 1 KXXXXXXXXXD 10

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73:\*

1: p1r1:\*

2: p1r2:\*

3: p1r3:\*

4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	73.1	284	2	T50261 conserved hypothet
2	19	73.1	378	2	E84806 probable elongatio
3	18	69.2	100	1	R3LV14 ribosomal protein
4	18	69.2	159	2	T48837 hypothetical prote
5	18	69.2	161	2	A82401 transcription regu
6	18	69.2	169	2	S07748 hypothetical prote
7	18	69.2	171	2	A96723 hypothetical prote
8	18	69.2	212	2	C90786 probable tet opero
9	18	69.2	212	2	A85646 probable transcrip
10	18	69.2	212	2	C64843 hypothetical prote
11	18	69.2	217	2	B64080 hypothetical prote
12	18	69.2	237	2	T14770 hypothetical prote
13	18	69.2	247	2	S01924 hypothetical prote
14	18	69.2	247	2	F37386 hypothetical prote
15	18	69.2	278	2	A83091 hypothetical prote
16	18	69.2	344	2	H88195 hypothetical prote
17	18	69.2	344	2	S22796 probable portal pr
18	18	69.2	365	2	JC3559 xylanase A - Stre
19	18	69.2	383	2	T34603 flagellin flhC-I -
20	18	69.2	413	2	S73643 MG349 homolog G12
21	18	69.2	427	2	A41790 actin like protein
22	18	69.2	477	1	US0589 endo-1,4-beta-xyla
23	18	69.2	543	2	T32496 hypothetical prote
24	18	69.2	837	2	S64785 hypothetical prote
25	18	69.2	858	2	T24062 hypothetical prote
26	18	69.2	863	2	G96964 probable permeal
27	18	69.2	1035	2	T16588 hypothetical prote
28	18	69.2	1042	2	T16169 hypothetical prote
29	18	69.2	1279	2	T17194 protoporphylin IX

30	17	65.4	109	2	A69448 hypothetical prote
31	17	65.4	118	2	T15655 hypothetical prote
32	17	65.4	137	2	G71541 probable comp-depe
33	17	65.4	137	2	C81695 cyclic nucleotide-
34	17	65.4	156	2	S23357 H+-transporting tw
35	17	65.4	185	2	G64084 3-methyladenine DN
36	17	65.4	186	2	S66070 conserved hypothet
37	17	65.4	207	2	AC2950 transcription regu
38	17	65.4	213	2	H98332 probable transcrip
39	17	65.4	217	1	B64222 heat shock protein
40	17	65.4	223	2	H84977 ribose-5-phosphate
41	17	65.4	226	2	A86268 hypothetical prote
42	17	65.4	234	1	XCSAS1 toxic shock syndro
43	17	65.4	234	2	B83992 toxic shock syndro
44	17	65.4	234	2	C69519 conserved hypothet
45	17	65.4	239	2	A83953 uridylylate kinase s

#### ALIGNMENTS

##### RESULT 1

T50261

conserved hypothetical protein SPAC890.05 with G-patch domain [Imported] - fission ye

C:Species: Schizosaccharomyces pombe

C>Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 09-Jun-2000

C/Accession: T50261

R:Saunders, D.; Harris, D.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, December 1999

A:Reference number: Z25052

A:Accession: T50261

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-284 <SNU>

A:Cross-references: EMBL:AL133498; P1DN:CAB63496.1; GSPDB:GN00066; SPDB:SPAC890.05

A:Experimental source: strain 972h(-); cosmid c890

C:Genetics:

A:Gene: SPDB:SPAC890.05

A:Map position: 1

##### Query Match

Best Local Similarity 73.1%; Score 19; DB 2; Length 284;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KXXXXXXXXXD 10

DB 179 KKTSTGSAID 188

##### RESULT 2

E84806 probable elongation factor [Imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C/Accession: E84806

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Beaulieu, M.L.; Town, C.D.; Fujii, C.Y.

M.; Koo, H.; Moffatt, K.S.; Cronin, L.A.; Shen, M.; Venkatesh, S.E.; Umayam, L.; Taiton,

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter

Nature 402, 761-768, 1999

A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: AB4420; M0ID:20083487; PMID:10617197

A:Accession: E84806

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-378 <STO>

A:Cross-references: GB:AE002093; NID:g3786016; P1DN:AAC67362.1; GSPDB:GN00139

C:Genetics:

A:Gene: AC2938560

A:Map position: 2

##### Query Match

Best Local Similarity 73.1%; Score 19; DB 2; Length 378;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy	1	KKKXXXXXXLD	10
Db	13	KKAADAAALD	22

### RESULT 3

ribosomal protein S14, chloroplast - liverwort (*Marchantia polymorpha*) chloroplast  
C:Species: chloroplast *Marchantia polymorpha*  
C:Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 22-Jun-1999  
C:Accession: A02733; S01876

submitted to the EMBL Data Library, October 1986  
A:Reference number: A00150  
A:Accession: A02733  
A:Molecule type: DNA  
A:Residues: 1-100 <OH>  
R:Ohyama, K.; Fukuzawa, H.; Kohchi, T.; Shirai, H.; Sano, S.; Umesono, K.; Shirai, H.  
Nature 322, 572-574, 1986  
A>Title: Chloroplast gene organization deduced from complete sequence of liverwort Marchantia polymorpha chloroplast genome. II. Genomic organization of the rbcL gene  
A:Reference number: A38014  
A:Contents: annotation; gene organization, sites, features  
R:Umesono, K.; Inokuchi, H.; Shikata, Y.; Takeuchi, M.; Chang, Z.; Fukuzawa, H.; Kohchi, T.; Mori, M.  
J. Mol. Biol. 203, 299-331, 1988  
A>Title: Structure and organization of Marchantia polymorpha chloroplast genome. II. Genomic organization of the rbcL gene  
A:Reference number: S01567; MUID:89068686; PMID:2974065  
A:Accession: S01876  
A:Molecule type: DNA  
A:Residues: 1-100 <DME>  
A:Cross-references: GB:X0445; GB:IY0086; MID:g11640; PIDN:CAA28083.1; PID:g11670  
A:Genetics:

C;Superfamily: Escherichia coli ribosomal protein S14  
C;Keywords: chloroplast; protein biosynthesis; ribosome

```

Qy      1 KXXXXXXXXLD 10
         ||      ||
Db      29 KKITETSSLD 38

```

## RESULT 4

hypothetical protein 6882.210 [imported] -Neurospora crassa  
C:Species: Neurospora crassa  
C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 16-Feb-2001  
C:Accession: T46887  
R:Schulze, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura  
submitted to the Protein Sequence Database, April 2000  
;Reference number: 224541

C: Superfamily: Neurospora crassa hypothetical protein 68B2.210

Query Match	69.2%	Score 18;	DB 2;	Length 159;
Best Local Similarity	40.0%	Pred. No. 12;		
Matches	4;	Conservative	0;	Mismatches 6;
			Indels 0;	Gaps 0;

```

QY      1 KXXXXXXXXLD 10
      ||      ||
Db      71 KKTALQSLD 80

```

RESULT 5  
A82401  
transcription regulator Tatp family [imported] - With 4 cholesterol /sterol NMR

Query Match	69.28;	Score 18;	DB 2;	length 161;
Best Local Similarity	40.0%;	Pred. No. 12;		
Matches	4;	Conservative	0;	Mismatches 6;
				Indels 0;
				Gaps 0;

Qy	1	KKKXXXXXXLD	10
Db	62	KKTFASALD	71

## RESULT

hypothetical protein 13 - Parametium tetraurelia mitochondrion  
C:Species: mitochondrion Parametium tetraurelia  
C:Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 07-Dec-1999  
C:Accession: S07748  
R:Pritchard, A.E.; Sellhammer, J.J.; Mahalingam, R.; Sadle, C.L.; Vennit, S.E.; Cummins  
Nucleic Acids Res. 18, 173-180, 1990  
A:Title: Nucleotide sequence of the mitochondrial genome of Parametium.  
A:Reference number: S07725; MUID:90174913; PMID:2308823

Query Match	69.28;	Score 18;	DB 2;	Length 169;
Best Local Similarity	40.08;	Pred. No. 13;		
Matches 4;	Conservative 0;	Mismatches 6;	Indels 0;	Gaps 0

```
QY      1 KKKXXXXXXLD 10
        ||      ||
Db      36 KKTSTYSKLD 45
```

## RESULT 7

hypothetical protein F20P5\_23 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: A96723  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federlspital, N.A.; Kaul, S.; White, O.; Alon  
Chung, C.W.; M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,  
ansen, N.F.; Hughes, B.; Hutzar, L.

Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marshall,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: A96723  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-171 <STO>  
A:Cross-References: GB:AE005173; NID:g2194134; PIDN:AAB61109.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: F20P5.23  
A:Map position: 1

Query Match 69.2%; Score 18; DB 2; Length 171;  
Best Local Similarity 40.0%; Pred. No. 13;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXLD 10  
DB 16 KKLTTTDL 25

RESULT 8  
C90786  
Probable tet operon regulator [Imported] - *Escherichia coli* (strain O157:H7, substrain F  
C:Species: *Escherichia coli*  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
C:Accession: C90786  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gene  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: C90786  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-212 <HAY>  
A:Cross-References: GB:BA000007; PIDN:BBB34682.1; PID:g13360719; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RMD 0509952  
C:Genetics:  
A:Gene: ECS1259

Query Match 69.2%; Score 18; DB 2; Length 212;  
Best Local Similarity 40.0%; Pred. No. 15;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXLD 10  
DB 20 KKAIIISALD 29

RESULT 9  
A85646  
Probable tet operon regulator ydcC [Imported] - *Escherichia coli* (strain O157:H7, substrain  
C:Species: *Escherichia coli*  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: A85646  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206531  
A:Accession: A85646  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-212 <STO>  
A:Cross-References: GB:AE005174; NID:g12514373; PIDN:AAG55629.1; GSPDB:GN00145; UNCP:215  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:

A:Gene: ydcC

Query Match 69.2%; Score 18; DB 2; Length 212;  
Best Local Similarity 40.0%; Pred. No. 15;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXLD 10  
DB 20 KKAIIISALD 29

RESULT 10  
C64843  
Probable transcription regulator ydcC - *Escherichia coli* (strain K-12)  
C:Species: *Escherichia coli*  
C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002  
C:Accession: C64843  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of *Escherichia coli* K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503  
A:Accession: C64843  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-212 <BLAT>  
A:Cross-References: GB:AE000203; GB:U00096; NID:g1787248; PIDN:AAC74098.1; PID:g17872  
A:Experimental source: strain K-12, substrain MG1655  
C:Genetics:  
A:Gene: ydcC  
C:Keywords: DNA binding; transcription regulation  
F:39-58/Region: helix-turn-helix motif

Query Match 69.2%; Score 18; DB 2; Length 212;  
Best Local Similarity 40.0%; Pred. No. 15;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXLD 10  
DB 20 KKAIIISALD 29

RESULT 11  
B64080  
Hypothetical protein HI0601 - *Haemophilus influenzae* (strain Rd KW20)  
C:Species: *Haemophilus influenzae*  
C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 08-Oct-1999  
C:Accession: B64080  
R:Eisenschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kirlavage  
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Feldman  
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.  
Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter  
A:Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.  
A:Reference number: A64000; MUID:95350630; PMID:7542800  
A:Accession: B64080  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-217 <YIG>  
A:Cross-References: GB:U32741; GB:L42023; NID:g1573582; PIDN:AAC22258.1; PID:g1573592

Query Match 69.2%; Score 18; DB 2; Length 217;  
Best Local Similarity 40.0%; Pred. No. 16;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXLD 10  
DB 161 KKGSGATLD 170

RESULT 12  
T14770  
Hypothetical protein DKFZP566E144.1 - human

C:Species: Homo sapiens (man)  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 11-Jan-2000  
 C:Accession: F37386  
 R:Biocheck: H.; Boecher, M.; Brandt, P.; Nemes, H.W.; Gassenhuber, J.; Wiemann, S.  
 submitted to the Protein Sequence Database, August 1999  
 A:Reference number: 218182  
 A:Accession: F14770  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-237 <BILO>  
 A:Cross-references: EMBL:AL110239  
 A:Experimental source: fetal kidney; clone DKFZp566E144  
 C:Genetics:  
 A:Note: DKFZp566E144.1  
 C:Superfamily: human conserved hypothetical protein DKFZp566E144.1

Query Match 69.2% Score 18; DB 2; Length 237;  
 Best Local Similarity 40.0%; Pred. No. 17;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXLD 10  
 || ||  
 Db 189 KKAASHRALD 198

RESULT 13  
 S01924  
 hypothetical protein 6 - Chlamydia trachomatis plasmids  
 N:Alternate names: hypothetical protein P-10  
 C:Species: Chlamydia trachomatis  
 C:Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 08-Oct-1999  
 C:Accession: S01924; S00794  
 R:Command: M.; Ricci, S.; Ratti, G.  
 M01. Microbiol. 2, 531-538, 1988  
 A:Title: The structure of a plasmid of Chlamydia trachomatis believed to be required for  
 A:Reference number: S01920; MUID:89013895; PMID:2845228  
 A:Accession: S01924  
 A:Molecule type: DNA  
 A:Residues: 1-247 <CON>  
 A:Cross-references: EMBL:X07547; NID:940730; PIDN:CAA30426.1; PID:9581014  
 A:Experimental source: plasmid pCHL1  
 A:Note: the authors translated the initiation codon GTG for residue 1 as Val  
 R:Hall, C.; Ward, M.E.; Clarke, I.N.  
 Nucleic Acids Res. 16, 4053-4067, 1988  
 A:Title: Analysis of the entire nucleotide sequence of the cryptic plasmid of Chlamydia  
 A:Reference number: S00788; MUID:88233998; PMID:2836808  
 A:Accession: S00794  
 A:Molecule type: DNA  
 A:Residues: 175-247 <HAT>  
 A:Cross-references: EMBL:X06707  
 A:Experimental source: plasmid pLGV440  
 C:Genetics:  
 A:Genome: Plasmid  
 A:Start codon: GTG

Query Match 69.2% Score 18; DB 2; Length 247;  
 Best Local Similarity 40.0%; Pred. No. 18;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXLD 10  
 || ||  
 Db 13 KKNQTAASLD 22

RESULT 14  
 F37386  
 hypothetical protein 6 - Chlamydia trachomatis (serotype D) plasmid pCHL1  
 C:Species: Chlamydia trachomatis  
 C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 17-Feb-1994  
 C:Accession: F37386  
 R:Command: M.; Ricci, S.; Cevenini, R.; Ratti, G.  
 Plasmid 23, 149-154, 1990  
 A:Title: Diversity of the Chlamydia trachomatis common plasmid in biovars with different

A:Reference number: A37386; MUID:90301796; PMID:2194229  
 A:Accession: F37386  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-247 <CON>  
 A:Cross-references: GB:J03321  
 C:Genetics:  
 A:Genome: plasmid

Query Match 69.2% Score 18; DB 2; Length 247;  
 Best Local Similarity 40.0%; Pred. No. 18;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXLD 10  
 || ||  
 Db 13 KKNQTAASLD 22

RESULT 15  
 AB3091  
 hypothetical protein Atu4350 [imported] - Agrobacterium tumefaciens (strain C58, Dupo  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
 C:Accession: AB3091  
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo  
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavina, T.; Levy, R.; Li, M.; MCI  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Tso, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam  
 ster, E.W.  
 A:Title: The genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A:Reference number: AB2577; PMID:11743193  
 A:Accession: AB3091  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-278 <KUR>  
 A:Cross-references: GB:AE006689; PIDN:AAL45144.1; PID:917742817; GSPDB:GN00187  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: Atu4350  
 A:Map position: linear chromosome

Query Match 69.2% Score 18; DB 2; Length 278;  
 Best Local Similarity 40.0%; Pred. No. 20;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXLD 10  
 || ||  
 Db 184 KKLAAFTFLD 193

Search completed: June 5, 2003, 15:33:59  
 Job time : 11.6485 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:34:07 ; Search time 6.51515 Seconds

(without alignments)  
63.661 Million cell updates/sec

Title: US-09-150-947F-14

Sequence: 1 KKKXXXXXXD 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	69.2	100	1	RR14_MARPO
2	18	69.2	152	1	NIJ1_HUMAN
3	18	69.2	152	1	NIJ1_MOUSE
4	18	69.2	152	1	NIJ1_RAT
5	18	69.2	169	1	YMI3_PASTE
6	18	69.2	212	1	YCDC_ECOLI
7	18	69.2	217	1	TFOX_HAEIN
8	18	69.2	237	1	ORN_HUMAN
9	18	69.2	247	1	GPED_CHLTR
10	18	69.2	344	1	VPO_BPP2
11	18	69.2	365	1	FLC1_PROMI
12	18	69.2	413	1	Y349_MYCPN
13	18	69.2	427	1	ARE3_SCHPO
14	18	69.2	477	1	XVNA_STRLI
15	18	69.2	477	1	YL34_YEAST
16	18	69.2	837	1	BCRH_CHLVI
17	18	69.2	1279	1	EX7S_PASMO
18	18	69.2	118	1	YQ93_CAEEL
19	18	69.2	142	1	NIJ2_HUMAN
20	18	69.2	143	1	NIJ2_MOUSE
21	18	69.2	144	1	NIJ2_RAT
22	18	69.2	147	1	ATP5_MYCGA
23	18	69.2	156	1	ATP5_ODOSI
24	18	69.2	185	1	3MGA_HAEIN
25	18	69.2	186	1	YABF_BACST
26	18	69.2	211	1	ORN_DROME
27	18	69.2	217	1	GRPE_MYCGE
28	18	69.2	223	1	RP1A_BUCAL
29	18	69.2	234	1	YB32_MYCPN
30	18	69.2	256	1	ETC1_MYCPN
31	18	69.2	266	1	ETC2_STRAU
32	18	69.2	266	1	ETC3_STRAU
33	18	69.2	266	1	ETC3_STRAU

34	17	65.4	272	1	TRUB_CAMJE	09p12 campylobact
35	17	65.4	276	1	MYO1_ONCMY	091205 oncohyacin
36	17	65.4	307	1	ERA_NEIMA	091v42 neisera m
37	17	65.4	307	1	ERA_NEIMA	091v42 neisera m
38	17	65.4	347	1	Y999_MYCPN	091v42 neisera m
39	17	65.4	363	1	LE3A_ASPEG	091v42 neisera m
40	17	65.4	368	1	DXR_HELPJ	091v42 neisera m
41	17	65.4	390	1	CNX2_ARATH	091v42 neisera m
42	17	65.4	399	1	Y199_ARCFU	091v42 neisera m
43	17	65.4	413	1	YB44_MYCPN	091v42 neisera m
44	17	65.4	457	1	PRIC_STRGR	091v42 neisera m
45	17	65.4	467	1	LEU2_ANASP	08yx02 anabaena sp

## ALIGNMENTS

RESULT 1  
ID RR14\_MARPO STANDARD: PRT: 100 AA.  
AC P06370:  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Chloroplast 30S ribosomal protein S14.  
GN RPS14.  
OS Marchantia polymorpha (Liverwort).  
OC Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Marchantiophyta;  
OC Marchantiopsida; Marchantiales; Marchantaceae; Marchantia.  
OX NCBI\_TaxID=3197;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85087956; PubMed=6393057;  
RA Umesono K., Inokuchi H., Ohyama K., Ozeki H.;  
RT "Nucleotide sequence of Marchantia polymorpha chloroplast DNA: a  
RT region possibly encoding three tRNAs and three proteins including a  
RT homologue of E. coli ribosomal protein S14."  
RL Nucleic Acids Res. 12:9551-9565(1984).  
[2]  
RP SEQUENCE FROM N.A.  
RA Ohyama K., Fukuzawa H., Kohchi T., Shirai H., Sano T., Sano S.,  
RA Umesono K., Shiki Y., Takeuchi M., Chang Z., Aota S., Inokuchi H.,  
RA Ozeki H.;  
RT "Chloroplast gene organization deduced from complete sequence of  
RT liverwort Marchantia polymorpha chloroplast DNA."  
RL Nature 322:572-574(1986).  
-1- SIMILARITY: BELONGS TO THE S14P FAMILY OF RIBOSOMAL PROTEINS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
DR EMBL: X01647; CA25805.1; -  
DR EMBL: X04465; CA28083.1; -  
DR PIR: A02733; R3LV14.  
DR InterPro: IPR001209; Ribosomal\_S14.  
DR Pfam: PF00253; Ribosomal\_S14; 1.  
DR PROSITE: PS00527; RIBOSOMAL\_S14; 1.  
KW Ribosomal protein; chloroplast.  
SQ SEQUENCE 100 AA; 11880 MW; 5AB14F480963EEB0 CRC64;  
Query Match 69.2%; Score 18; DB 1; Length 100;  
Best Local Similarity 40.0%; Pred. No. 4.4;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
1 KKKXXXXXXD 10  
||  
||

DB 29 KKITSSLD 38

RESULT 2

NIJL\_HUMAN STANDARD: PRT: 152 AA.

AC Q92982; Q92982; 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Nijurin 1 (Nerve injury-induced protein 1).

GN NINJ1

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96374367; PubMed=8780658;

RA Araki T., Milbrandt J.;

RT "Ninjurin, a novel adhesion molecule, is induced by nerve injury and promotes axonal growth."

RL Neuron 17:353-361(1996).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=98126436; PubMed=9465296;

RA Chadwick B.P., Heath S.K., Williamson J., Obermayr F., Patel L.,

RA Sheer D., Frischauf A.M.;

RT "The human homologue of the ninjurin gene maps to the candidate region of hereditary sensory neuropathy type I (HSNI)."

RL Genomics 47:58-63(1998).

RN [3]

RP SEQUENCE FROM N.A.

RX TISSUE=Lung;

RC Strausberg R.;

RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

RN [4]

RP CHARACTERIZATION.

RX MEDLINE=97407927; PubMed=9261151;

RA Araki T., Zimonjic D.B., Popescu N.C., Milbrandt J.;

RT "Mechanism of homophilic binding mediated by ninjurin, a novel widely expressed adhesion molecule."

RL J. Biol. Chem. 272:21373-21380(1997)

CC -1- FUNCTION: HOMOPHILIC CELL ADHESION MOLECULE THAT PROMOTES AXONAL GROWTH. MAY PLAY A ROLE IN NERVE REGENERATION AND IN THE FORMATION AND FUNCTION OF OTHER TISSUES. CELL ADHESION REQUIRES DIVALENT CATIONS.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED IN BOTH ADULT AND EMBRYONIC TISSUES, PRIMARILY THOSE OF EPITHELIAL ORIGIN.

CC -1- INDUCTION: BY NERVE INJURY BOTH IN DORSAL ROOT GANGLION NEURONS AND IN SCHWANN CELLS.

CC -1- SIMILARITY: BELONGS TO THE NINJURIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

CC EMBL; U72661; AAB17560.1; -

DR EMBL; U91512; AAC14593.1; -

DR EMBL; AF029251; AAC39574.1; -

DR EMBL; BC004440; AA044440.1; -

DR Genew: HGNC:7824; NINJ1.

DR MIM; 602062; -

DR Cell adhesion; Transmembrane.

FT DOMAIN 1 80 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 81 101 POTENTIAL.

FT DOMAIN 102 120 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 121 141 POTENTIAL.

FT DOMAIN 142 152 EXTRACELLULAR (POTENTIAL).

FT CONFLICT 110 110 D -> A (IN REF. 3).

SO SEQUENCE 152 AA; 16389 MW; FEACAD8001597855 CRC64;

Query Match

Best Local Similarity 40.0%; Score 18; DB 1; Length 152;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

DB 44 KKSAAESMLD 53

QY 1 KXXXXXXLD 10

DB 11 11

DB 44 KKSAAESMLD 53

RESULT 3

NIJL\_MOUSE STANDARD: PRT: 152 AA.

AC Q70131;

DT 30-MAY-2000 (Rel. 39, Created)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Ninjurin 1 (Nerve injury-induced protein 1).

GN NINJ1

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98126436; PubMed=9465296;

RA Chadwick B.P., Heath S.K., Williamson J., Obermayr F., Patel L.,

RA Sheer D., Frischauf A.M.;

RT "The human homologue of the ninjurin gene maps to the candidate region of hereditary sensory neuropathy type I (HSNI)."

RL Genomics 47:58-63(1998).

RN [2]

RP SEQUENCE FROM N.A.

RX Moon A.R., Kim J.W., Hong Y.M., Oh G.T., Chang S.Y., Lee K.S.,

RA Choe I.S.;

RT "Mus musculus ninjurin genomic DNA."

RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: HOMOPHILIC CELL ADHESION MOLECULE THAT PROMOTES AXONAL GROWTH. MAY PLAY A ROLE IN NERVE REGENERATION AND IN THE FORMATION AND FUNCTION OF OTHER TISSUES. CELL ADHESION REQUIRES DIVALENT CATIONS (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- INDUCTION: BY NERVE INJURY.

CC -1- SIMILARITY: BELONGS TO THE NINJURIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

CC EMBL; U91513; AAC14594.1; -

DR EMBL; AF219626; AAC32161.1; -

DR MGD; MGI:1196617; NINJ1.

DR Cell adhesion; Transmembrane.

FT DOMAIN 1 79 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 80 100 POTENTIAL.

FT DOMAIN 101 120 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 121 141 POTENTIAL.

FT DOMAIN 142 152 EXTRACELLULAR (POTENTIAL).

SO SEQUENCE 152 AA; 16555 MW; E261CB447BC0A286 CRC64;

Query Match

Best Local Similarity 69.2%; Score 18; DB 1; Length 152;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXXXXXXLD 10

DB 11 11

DB 44 KKSAAESMD 53

RESULT 4  
NINJ1\_RAT STANDARD: PRT; 152 AA.

AC P70617;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DE 16-OCT-2001 (Rel. 40, Last annotation update)  
GN NINJ1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96374367; PubMed=8780658;  
RA Araki T., Milbrandt J.;  
RT "Ninjurin, a novel adhesion molecule, is induced by nerve injury and promotes axonal growth."  
RL Neuron 17:353-361(1996).  
CC -1- FUNCTION: HOMOPHILIC CELL ADHESION MOLECULE THAT PROMOTES AXONAL GROWTH. MAY PLAY A ROLE IN NERVE REGENERATION AND IN THE FORMATION AND FUNCTION OF OTHER TISSUES. CELL ADHESION REQUIRES DIVALENT CATIONS (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- SIMILARITY: BELONGS TO THE NINJURIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U72660; AAB17559.1; -.  
KW Cell adhesion; Transmembrane.  
FT DOMAIN 1 79 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 80 100 POTENTIAL.  
FT DOMAIN 101 120 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 121 141 POTENTIAL.  
FT DOMAIN 142 152 EXTRACELLULAR (POTENTIAL).  
SQ SEQUENCE 152 AA; 16539 MW; 414B3BDE1807E80 CRC64;  
Query Match 69.2%; Score 18; DB 1; Length 152;  
Best Local Similarity 40.0%; Pred. No. 6.5;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXID 10  
DB 44 KKSAAESMD 53

RESULT 5  
YMI3\_PART2 STANDARD: PRT; 169 AA.

AC P15614;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DE 01-APR-1990 (Rel. 14, Last annotation update)  
GN Hypothetical 20.2 kDa protein (ORF13).  
OS Paramoecium tetraurelia.  
OC Mitochondrion.  
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Penicillata;  
OC Paramoecium.  
OX NCBI\_TaxID=5888;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Stock 51;

RX MEDLINE=90174913; PubMed=2308823;  
RA Pritchard A.E., Sellhammer J.J., Mahalingam R., Sable C.L.,  
RA Venuti S.E., Cummings D.J.;  
RT "Nucleotide sequence of the mitochondrial genome of Paramoecium."  
RL Nucleic Acids Res. 18:1173-180(1990).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X15917; CAA34057.1; -.  
DR PIR: S07748; S07748.  
KW Hypothetical protein; Mitochondrion.  
SQ SEQUENCE 169 AA; 20303 MW; D991AD99A8EBA174 CRC64;  
Query Match 69.2%; Score 18; DB 1; Length 169;  
Best Local Similarity 40.0%; Pred. No. 7.2;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXID 10  
DB 36 KKTSTYSKID 45

RESULT 6  
YCDC\_ECOLI STANDARD: PRT; 212 AA.

AC P75899;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DE 16-OCT-2001 (Rel. 40, Last annotation update)  
GN YCDC OR B1013.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12."  
RL Science 277:1234-1238(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12;  
RX MEDLINE=97061202; PubMed=8905232;  
RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,  
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,  
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,  
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,  
RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,  
RA Yano M., Horikuchi T.;  
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 12.7-28.0 min region on the linkage map."  
RL DNA Res. 3:137-155(1996).  
CC -1- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTOMAL REGULATORS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

or send an email to [license@lsb.slb.ch](mailto:license@lsb.slb.ch)).

CC EMBL; AE000203; AAC74098.1; -  
 DR EMBL; D90738; BAA35790.1; -  
 DR Ecogene; EG12301; YCOC.  
 DR InterPro; IPR001647; HTH\_Tetr.  
 DR Pfam; PF00440; tetr.1.  
 DR PROSITE; PS01081; HTH\_TETR\_FAMILY; FALSE\_NEG.  
 KW Hypothetical protein; Transcription regulation; DNA-binding;  
 KM Complete proteome.  
 FT DNA\_BIND 39 58 H-T-H MOTIF (POTENTIAL).  
 SQ SEQUENCE 212 AA: 23687 MW: E245985DFAC277A CRC64;

Query Match 69.2%; Score 18; DB 1; Length 212;  
 Best Local Similarity 40.0%; Pred. No. 9;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKKXXXXXLD 10  
 DB 20 KKAISALD 29

RESULT 7  
 TFOX\_HAEIN STANDARD; PRT: 217 AA.  
 AC P43779;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE DNA transformation protein tfox (Competence activator) (Protein sxy).  
 GN TFOX OR SXY OR HT0601.  
 OS Haemophilus influenzae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OX NCBI\_TaxID=727;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Rd;  
 RX MEDLINE=95241551; PubMed=7724607;  
 RA Zully J.J., Barcak G.J.;  
 RT "Identification of a DNA transformation gene required for com101A+  
 RT expression and supertransformer phenotype in Haemophilus  
 RT influenzae";  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:3616-3620(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Rd / KW20 / ATCC 51907;  
 RX MEDLINE=95350630; PubMed=7542800;  
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Kesteven A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudak D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghagen N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae  
 RT Rd.";  
 RL Science 269:496-512(1995).  
 RN [3]  
 RP IDENTIFICATION OF TFOX AS SXY AND VARIANT SXY-1.  
 RX MEDLINE=95050244; PubMed=7961436;  
 RA Williams P.M., Bannister L.A., Redfield R.J.;  
 RT "The Haemophilus influenzae sxy-1 mutation is in a newly identified  
 RT gene essential for competence";  
 RL J. Bacteriol. 176:6789-6794(1994).  
 CC -1- FUNCTION: REQUIRED FOR DNA TRANSFORMATION. POSITIVELY REGULATES  
 CC GENE REQUIRED FOR DNA TRANSFORMATION (LATE COMPETENCE-SPECIFIC  
 CC GENES).  
 CC -1- SIMILARITY: TO E.COLI YCCR.  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.lsb.slb.ch/announce/>  
 CC or send an email to [license@lsb.slb.ch](mailto:license@lsb.slb.ch)).

CC EMBL; U13205; AAC43320.1; -  
 DR EMBL; U13205; AAC43321.1; -  
 DR EMBL; U32741; AAC22258.1; -  
 DR TIGR; HT0601; -  
 KW Competence; Transcription regulation; Activator;  
 KM Alternative initiation; Complete proteome.  
 FT CHAIN 1 217 DNA TRANSFORMATION PROTEIN TFOX.  
 FT CHAIN 40 217 DNA TRANSFORMATION PROTEIN TFOX-SHORT.  
 FT INIT MET 40 40 FOR TFOX-SHORT.  
 FT VARIANT 19 19 V -> I (IN SXY-1; CAUSES A 100 X TO 1000  
 X INCREASE IN SPONTANEOUS NATURAL  
 COMPETENCE).  
 FT CONFLICT 203 203 W -> L (IN REF. 1).  
 SQ SEQUENCE 217 AA: 25074 MW: 4A6E30E64008CC3F CRC64;

Query Match 69.2%; Score 18; DB 1; Length 217;  
 Best Local Similarity 40.0%; Pred. No. 9.2;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKKXXXXXLD 10  
 DB 161 KKSQSGATLD 170

RESULT 8  
 ORN\_HUMAN STANDARD; PRT: 237 AA.  
 AC Q9Y3B8; Q9UFY7;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Oligoribonuclease, mitochondrial precursor (EC 3.1.1.-) (Small  
 DE fragment nuclease) (CGI-114).  
 GN SFN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (SHORT ISOPFORM).  
 RC MEDLINE=20272150; PubMed=10810093;  
 RA Lai C.-H., Chou C.-Y., Chang L.-Y., Liu C.-S., Lin W.-C.;  
 RT "Identification of novel human genes evolutionarily conserved in  
 RT Caenorhabditis elegans by comparative proteomics";  
 RL Genome Res. 10:703-713(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A. (LONG ISOPFORM).  
 RC TISSUE=Kidney.  
 RX MEDLINE=21154917; PubMed=11230166;  
 RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,  
 RA Ansorge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,  
 RA Lauber J., Duesterhoeft A., Beyer A., Kohrer K., Strack N.,  
 RA Mewes H.-W., Ottenwaelder B., Obermaler B., Tampe J., Heubner D.,  
 RA Wandut R., Korn B., Klein M., Poustka A.;  
 RT "Towards a catalog of human genes and proteins: sequencing and  
 RT analysis of 500 novel complete protein coding human cDNAs";  
 RL Genome Res. 11:422-435(2001).  
 RN [3]  
 RP CHARACTERIZATION.  
 RX MEDLINE=20408924; PubMed=10851236;  
 RA Nguyen L.H., Erzberger J.P., Root D.M., III;  
 RT "The human homolog of Escherichia coli Orn degrades small  
 RT single-stranded RNA and DNA oligomers";  
 RL J. Biol. Chem. 275:25900-25906(2000).  
 CC -1- FUNCTION: 3'-TO-5' EXORIBONUCLEASE SPECIFIC FOR SMALL  
 CC OLIGORIBONUCLEOTIDES. CC ACTIVE ON SMALL (PRIMARYLY </=5



CC NUCLEOTIDES IN LENGTH) SINGLE-STRANDED RNA AND DNA OLIGOMERS. MAY  
 CC HAVE A ROLE FOR CELLULAR NUCLEOTIDE RECYCLING.  
 CC -1- COFACTOR: MANGANESE.  
 CC -1- SUBUNIT: HOMOTETRAMER (PROBABLE).  
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL (LONG ISOFORM) AND NUCLEAR  
 CC (SHORT ISOFORM).  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM/SPN-ALPHA (SHOWN  
 CC HERE) AND A SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: BELONGS TO THE OLIGORIBONUCLEASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF151872; AAD34109.1; -  
 DR EMBL; AL110239; CAB53690.1; -  
 DR InterPro: IPR000520; Exonuclease.  
 DR Pfam: PF00929; Exonuclease; 1.  
 DR SMART; SM00479; EXOIII; 1.  
 KW Hydroxylase; Exonuclease; Nuclease; Mitochondrion; Transit peptide;  
 KW Nuclear protein; Manganese; Alternative splicing.  
 FT TRANSIT 1  
 FT CHAIN ? 237  
 FT ACT SITE 164 164 POTENTIAL.  
 FT VARSPIC 1 38 MISSING (IN ISOFORM 2).  
 FT MORGAN 168 168 D->A: ABOLISHES ACTIVITY.  
 FT CONFLICT 103 103 R -> K (IN REF. 1).  
 SQ SEQUENCE 237 AA; 26861 MW; BF093A89078C13BF CRC64;

Query Match 69.2%; Score 18; DB 1; Length 237;  
 Best Local Similarity 40.0%; Pred. No. 10;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXD 10  
 DB 189 KKAASHRALD 198

RESULT 9  
 GPE6.CHLTR  
 ID GPE6.CHLTR STANDARD; PRT; 247 AA.  
 AC P10560; P08787; Q66433;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-JUL-1989 (Rel. 11, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Virulence plasmid protein pGP6-D (Protein P-10).  
 DE Chlamydia trachomatis.  
 OS Plasmid pLGVA40, Plasmid pCHL1, and Plasmid pCTT1.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_Taxid=813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-L2/434/Bu; PLASMID-pLGVA40;  
 RX MEDLINE-89013895; PubMed-2845228;  
 RA Comanducci M., Ricci S., Ratti G.,  
 RT "The structure of a plasmid of Chlamydia trachomatis believed to be  
 RT required for growth within mammalian cells."  
 RL Mol. Microbiol. 2:531-538(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-L1/440/LN; PLASMID-pLGVA40;  
 RX MEDLINE-88233998; PubMed-2836808;  
 RA Hatt C., Ward M.E., Clarke I.N.,  
 RT "Analysis of the entire nucleotide sequence of the cryptic plasmid of  
 RT Chlamydia trachomatis serovar L1. Evidence for involvement in DNA  
 RT replication."  
 RL Nucleic Acids Res. 16:4053-4067(1988).  
 RN [3]  
 RP SEQUENCE FROM N.A.

RC STRAIN-G0/86 / Serotype D; PLASMID-pCHL1;  
 RX MEDLINE-90301796; PubMed-2194229;  
 RA Comanducci M., Ricci S., Cevenini R., Ratti G.;  
 RT "Diversity of the Chlamydia trachomatis common plasmid in biovars  
 RT with different pathogenicity."  
 RL Plasmid 23:149-154(1990).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Serotype B; PLASMID-pCTT1;  
 RX MEDLINE-88177106; PubMed-3444859;  
 RA Stripakash K.S., Macavoy E.S.,  
 RT "Characterization and sequence of a plasmid from the trachoma biovar  
 RT of Chlamydia trachomatis."  
 RL Plasmid 18:205-214(1987).  
 CC -1- MISCELLANEOUS: pGP6-D IS REQUIRED FOR GROWTH WITHIN MAMMALIAN  
 CC CELLS.  
 CC -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF PLASMID pLGVA40.  
 CC -1- SIMILARITY: BELONGS TO THE UPF0137 (pGP6-D) FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X07547; CAA30426.1; -  
 DR EMBL; X06707; CAA29897.1; -  
 DR EMBL; J03321; AAA91574.1; -  
 DR EMBL; M19487; AAB02591.1; -  
 DR PIR; S01924; S01924.  
 DR PIR; S00794; S00794.  
 DR InterPro: IPR005350; UPF0137.  
 DR Pfam: PF03677; UPF0137; 1.  
 KW Plasmid.  
 FT VARIANT 9 9 N -> D (IN PLASMIDS PCHL1 AND PCTT1).  
 FT VARIANT 47 47 Q -> R (IN PLASMIDS PCHL1 AND PCTT1).  
 FT VARIANT 183 186 VLDQ -> CMIE (IN PLASMID PCTT1).  
 FT VARIANT 194 194 K -> T (IN PLASMID PCTT1).  
 FT VARIANT 205 205 S -> L (IN PLASMID PCTT1).  
 FT VARIANT 222 222 M -> I (IN PLASMID PCTT1).  
 SQ SEQUENCE 247 AA; 28307 MW; F1FA11F29037C221 CRC64;

Query Match 69.2%; Score 18; DB 1; Length 247;  
 Best Local Similarity 40.0%; Pred. No. 10;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXD 10  
 DB 13 KKNQTAASLD 22

RESULT 10  
 VPO\_BPP2  
 ID VPO\_BPP2 STANDARD; PRT; 344 AA.  
 AC P25480;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Presumed portal vertex protein (GP0).  
 DE O.  
 OS Bacteriophage P2.  
 OC viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;  
 OC P2-like viruses.  
 OX NCBI\_Taxid=10679;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE-92115571; PubMed-1837355;  
 RA Linderoth N.A., Ziemann R., Hagggaard-Ljungquist E., Christie G.E.,  
 RA Calendar R.,  
 RT "Nucleotide sequence of the DNA packaging and capsid synthesis genes  
 RT of bacteriophage P2."

RL Nucleic Acids Res. 19:7207-7214(1991).  
 CC -1- FUNCTION: THE Q, P AND M PROTEINS ARE NEEDED TO PACKAGE DNA INTO  
 CC PROHEADS AND FOR THE CONVERSION OF PROHEADS TO CAPSIDS.  
 CC -1- SIMILARITY: STRONG, TO PHAGE HP1 PROTEIN ORF15 AND TO THE E. COLI  
 CC HYPOTHETICAL 36.8 kDa PROTEIN (ORF5) IN RETRON EC67.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF063097; AAD03268.1; --  
 DR PIR; S16410; S16410.  
 DR PIR; S22796; S22796.  
 KW Capsid assembly.  
 SQ SEQUENCE 344 AA; 39114 MW; FAD61987530029A7 CRC64;  
 Query Match 69.2%; Score 18; DB 1; Length 344;  
 Best Local Similarity 40.0%; Pred. No. 14;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 KKKXXXXXD 10  
 DB 270 KKASADLD 279  
 RESULT 11  
 FLCL\_PROMI STANDARD; PRT; 365 AA.  
 AC P42272;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Flagellin 1.  
 GN FLICL.  
 OS Proteus mirabilis.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 CC Proteus.  
 OX NCBI\_TaxID=584;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 1-29.  
 RC STRAIN-BB200;  
 RX MEDLINE=95011656; PubMed=7926835;  
 RA Belas R., Flaherty D.;  
 RT "Sequence and genetic analysis of multiple flagellin-encoding genes  
 RT from Proteus mirabilis.";  
 RL Gene 148:33-41(1994).  
 CC -1- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO  
 CC FORM THE FILAMENTS OF BACTERIAL FLAGELLA.  
 CC -1- INDUCTION: ALTHOUGH SWIMMER CELLS HAVE ONLY A FEW FLAGELLA, THE  
 CC ELONGATED SWARMER CELLS ARE PROPELLED COVERED BY THOUSANDS OF NEW  
 CC FLAGELLA SYNTHESIZED SPECIFICALLY IN RESPONSE TO GROWTH ON  
 CC SURFACES OR IN HIGHLY VISCOUS LIQUIDS.  
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF221596; AAB62396.1; --  
 DR InterPro; IPR001492; Flagellin.  
 DR InterPro; IPR001029; Flagellin\_C.  
 DR Pfam; PF00669; Flagellin\_N.1.  
 DR Pfam; PF00700; Flagellin\_C.1.  
 DR PRINTS; PR00207; Flagellin.  
 DR Pfam; PF000316; Flagellin\_C.1.

KW Flagella; Multigene family.  
 SQ SEQUENCE 365 AA; 39093 MW; 2A23106F040DB522 CRC64;  
 Query Match 69.2%; Score 18; DB 1; Length 365;  
 Best Local Similarity 40.0%; Pred. No. 15;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 KKKXXXXXD 10  
 DB 197 KKINAATLD 206  
 RESULT 12  
 Y349\_MYCPN STANDARD; PRT; 413 AA.  
 ID Y349\_MYCPN  
 AC P75253;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein MG349 homolog (G12\_orf413).  
 GN MPN525 OR MP317.  
 OS Mycoplasma pneumoniae.  
 CC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2104;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 29342 / M129;  
 RX MEDLINE=97105885; PubMed=8948633;  
 RA Himmelfeich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,  
 RA Herrmann R.;  
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
 RT pneumoniae.";  
 RL Nucleic Acids Res. 24:4420-4449(1996).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE000029; AAB95965.1; --  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 413 AA; 48246 MW; AACFC2C9802316F23 CRC64;  
 Query Match 69.2%; Score 18; DB 1; Length 413;  
 Best Local Similarity 40.0%; Pred. No. 17;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 KKKXXXXXD 10  
 DB 10 KATSFSGLD 19  
 RESULT 13  
 ARP3\_SCHPO STANDARD; PRT; 427 AA.  
 ID ARP3\_SCHPO  
 AC P32390;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Actin-like protein 3.  
 GN ACT3 OR SPAC630.03.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 MEDLINE=92108083; PubMed=1729722;  
 RX Lees-Miller J.P., Henry G., Helfman D.M.;

\*Identification of act2, an essential gene in the fission yeast  
Schizosaccharomyces pombe that encodes a protein related to actin.  
[2]  
Proc. Natl. Acad. Sci. U.S.A. 89:80-83(1992).

SEQUENCE FROM N.A.

RC STRAIN-972; PubMed-11859360;  
MEDLINE-21848401; PubMed-11859360;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Sgouras K., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Church C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Stevens K.,  
RA Taylor K., Taylor R.G., Tiley A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volkert G., Aert R., Robben J., Grynolprez B.,  
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
RA Goffeau A., Cadiet E., Dreano S., Gloux S., Lelaure V., Mottler S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerruti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Useery D., Barrett B.G., Nurse P.,  
RA "The genome sequence of Schizosaccharomyces pombe."  
RL Nature 415:871-880(2002).

CC -1- FUNCTION: PART OF A COMPLEX IMPLICATED IN THE CONTROL OF ACTIN  
POLYMERIZATION IN CELLS (BY SIMILARITY). MAY BE INVOLVED IN  
CYTOKINESIS.

CC -1- SUBUNIT: BELONGS TO A COMPLEX COMPOSED OF ARP2, ARP3, P41-ARC,  
P34-ARC, P21-ARC, P20-ARC AND P16-ARC (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE ACTIN FAMILY. ARP3 SUBFAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

DR EMBL: M81068; NOT ANNOTATED CDS.  
DR EMBL: AL109832; CAB52725.1; -.  
DR PIR: A41790; A41790.  
DR InterPro: IPR004000; Actin\_like.  
DR Pfam: PF000022; actin; 1.  
DR SMART: SM00268; ACTIN; 1.  
DR PROSITE: PS01132; ACTIN\_ACT\_LIKE; 1.  
KW Structural protein; Cytoskeleton.  
SQ SEQUENCE 427 AA; 47373 MW; C7909FEFE544789B CRC64;

Query Match 69.2%; Score 18; DB 1; Length 427;  
Best Local Similarity 40.0%; Pred. No. 17;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKKKXXXXXLD 10  
DB 82 KKASAGSYLD 91

RESULT 14  
ID XYNL STRLI STANDARD; PRT; 477 AA.  
AC P26514; P96464;  
DT 01-AUG-1992 (rel. 23. Created)  
DT 30-MAY-2000 (rel. 39; Last sequence update)

DT 15-JUN-2002 (rel. 41; Last annotation update)  
DE Endo-1,4-beta-xylanase A precursor (EC 3.2.1.8) (Xylanase A)  
DE (1,4-beta-D-xylan xylanohydrolase A).  
GN XLANA.  
OS Streptomyces lividans.  
OC Bacteria; Actinobacteria; Actinobacteriales; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI Taxid=11916;  
RN [1]  
RN SEQUENCE FROM N.A., AND SEQUENCE OF 42-92.  
RC STRAIN-66 / 1326;  
RA MEDLINE-92077439; PubMed-1743521;  
RA Shareck F., Roy C., Yaguchi M., Morosoli R., Kluepfel D.;  
RT "Sequences of three genes specifying xylanases in Streptomyces  
lividans."  
RL Gene 107:75-82(1991).

RN [2]  
RN REVISIONS TO 20 AND 140-141.  
RC STRAIN-66 / 1326;  
RA Shareck F.;  
RN submitted (JUL-1998) to the EMBL/Genbank/DBJ databases.  
RN [3]  
RN X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 42-340.  
RA MEDLINE-94342232; PubMed-8063693;  
RA Derwenda U., Swenson L., Green R., Wei Y.Y., Morosoli R., Shareck F.,  
RA Kluepfel D., Derwenda L.S.;  
RT "Crystal structure, at 2.6-A resolution, of the Streptomyces lividans  
xylanase A, a member of the F family of beta-1,4-D-glycanases."  
RL J. Biol. Chem. 269:20811-20814(1994).

CC -1- FUNCTION: CONTRIBUTES TO HYDROLYSE HEMICELLULOSE, THE MAJOR  
COMPONENT OF PLANT CELL-WALLS. XLANA AND XLAN SEM TO ACT  
SEQUENTIALLY ON THE SUBSTRATE TO YIELD XYLOBIOSIDE AND XYLOSE  
AS CARBON SOURCES.

CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-xylosidic  
linkages in xylans.

CC -1- PATHWAY: xylan degradation.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY F (FAMILY 10 OF GLYCOSYL  
HYDROLASES).

CC -----

CC -1- SIMILARITY: CONTAINS 1 RICIN B-TYPE LECTIN DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

DR EMBL: M64551; AAC26525.1; -.  
DR PIR: JS0589; JS0589.  
DR PDB: 1XAS; 3I-MAY-95.  
DR InterPro: IPR001000; Glyco\_hydro\_10.  
DR InterPro: IPR000772; Ricin\_B\_lectin.  
DR Pfam: PF00331; Glyco\_hydro\_10; 1.  
DR Pfam: PF00652; Ricin\_B\_lectin; 3.  
DR PRINTS: PR00134; GLYHDRLASE10.  
DR SMART: SM00458; RICIN; 1.  
DR PROSITE: PS00591; GLYCOSYL\_HYDROL\_F10; 1.  
DR PROSITE: PS00231; RICIN\_B\_LLECTIN; 1.  
KW Xylan degradation; Hydrolase; Glycosidase; Signal; Lectin;  
KW 3D-structure.

KW SIGNAL 1 41  
FT CHAIN 42 477 ENDO-1,4-BETA-XYLANASE A.  
FT DOMAIN 361 477 RICIN B-TYPE LECTIN.  
FT ACT\_SITE 169 169 PROTON DONOR.  
FT ACT\_SITE 277 277 NUCLEOPHILE.  
SQ SEQUENCE 477 AA; 51162 MW; E14A7FE37BDC68CC CRC64;

Query Match 69.2%; Score 18; DB 1; Length 477;  
Best Local Similarity 40.0%; Pred. No. 19;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXLD 10  
 DB 330 KKAATTAVID 339

RESULT 15  
 YL34\_YEAST

YL34\_YEAST STANDARD: PRT; 837 AA.

AC Q07844;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Hypothetical 93.1 kDa protein YL034C.  
 GN YL034C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Duesterhoeft A., Floeth M., Heuss-Neitzel D., Hilbert H., Moestl D.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1 SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----

CC  
 DR EMBL: Z73139; CA97483.1; -  
 DR SGD: S0003957; YL034C.  
 DR InterPro: IPR003953; AAA\_ATPase.  
 DR InterPro: IPR003959; AAA\_ATPase\_cent.  
 DR InterPro: IPR003960; AAA\_sub.  
 DR Pfam: PF00004; AAA; 2.  
 DR SMART: SM00362; AAA; 2.  
 DR PROSITE: PS00674; AAA; 1.  
 KW Hypothetical protein; ATP-binding; Repeat.  
 FT NP\_BIND 246 253 ATP (POTENTIAL).  
 FT NP\_BIND 574 581 ATP (POTENTIAL).  
 SO SEQUENCE 837 AA; 93069 MW; 133528B9D0987103 CRC64;

Query Match 69.2%; Score 18; DB 1; Length 837;  
 Best Local Similarity 40.0%; Pred. No. 33;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXLD 10  
 DB 7 KKNSTLTSILD 16

Search completed: June 5, 2003, 15:56:33  
 Job time : 8.51515 secs



```

ID 092VH8 PRELIMINARY; PRT; 378 AA.
AC 092VH8;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Putative elongation factor.
GN ATG38560.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanden S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhagen G.P., Preuss D., Niernan W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005499; AAC67362.1;
DR HSSP; P23193; 1FEI.
DR InterPro; IPR001222; TFIIS.
DR InterPro; IPR003618; TFS2_centre.
DR InterPro; IPR003617; TFS2_N.
DR Pfam; PF01096; TFIIS; 1.
DR SMART; SM00510; TFS2M; 1.
DR SMART; SM00509; TFS2N; 1.
DR SMART; SM00440; znf_C2C2; 1.
DR PROSITE; PS00466; TFIIS; 1.
KM Elongation factor.
SQ SEQUENCE 378 AA; 41979 MW; 9A176DC31FA8E885 CRC64;

Query Match
Best Local Similarity 73.1%; Score 19; DB 10; Length 378;
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKKXXXXXLD 10
DB 13 KKAADAAALD 22

RESULT 3
08VXZ6 PRELIMINARY; PRT; 378 AA.
ID 08VXZ6;
AC 08VXZ6;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative elongation factor.
GN ATG38560.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,

```

```

RA Kamluya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shim P., Southwick A., Shinozaki K.,
RA Davis R.W., Ecker J.R., Theologis A.;
RT "Full length cDNA of gene At2g38560 (GI:15224901).";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY074322; AAL67018.1;
DR InterPro; IPR001222; TFIIS.
DR InterPro; IPR003618; TFS2_centre.
DR InterPro; IPR003617; TFS2_N.
DR Pfam; PF01096; TFIIS; 1.
DR SMART; SM00510; TFS2M; 1.
DR SMART; SM00509; TFS2N; 1.
DR SMART; SM00440; znf_C2C2; 1.
DR PROSITE; PS00466; TFIIS; UNKNOWN_1.
KM Elongation factor.
SQ SEQUENCE 378 AA; 42011 MW; 04E5E924FBD9CDB CRC64;

Query Match
Best Local Similarity 73.1%; Score 19; DB 10; Length 378;
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKKXXXXXLD 10
DB 13 KKAADAAALD 22

RESULT 4
037174 PRELIMINARY; PRT; 1365 AA.
ID 037174;
AC 037174;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Replicase.
OS Bamboo mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potexvirus.
OX NCBI_TaxID=35286;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S;
RA Liao J.T., Hsu Y.H.;
RT "Full-length sequence of infectious bamboo mosaic virus genomic RNA.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF018156; AAB70562.1;
DR InterPro; IPR001788; RNA_dep_RNAPol2.
DR InterPro; IPR000606; Viral_helicase1.
DR Pfam; PF00978; RNA_dep_RNAPol2; 1.
DR Pfam; PF01443; Viral_helicase1; 1.
SQ SEQUENCE 1365 AA; 155219 MW; CA54D8F5997F158A CRC64;

Query Match
Best Local Similarity 73.1%; Score 19; DB 12; Length 1365;
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKKXXXXXLD 10
DB 1302 KKTATAYALD 1311

RESULT 5
065005 PRELIMINARY; PRT; 1365 AA.
ID 065005;
AC 065005;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ORF1-155K.
OS Bamboo mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potexvirus.
OX NCBI_TaxID=35286;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN-BAMV-O;  
 RA Hsu Y.-H.;  
 RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BAMV-O;  
 RX MEDLINE-94358760; PubMed-8077956;  
 RA Lin N., Lin B., Lo N., Hu C., Chow T., Hsu Y.;  
 RT "Nucleotide sequence of the genomic RNA of bamboo mosaic potexvirus.";  
 RL J. Gen. Virol. 75:2513-2518(1994).  
 DR EMBL: D26017; BAA05033.1; -;  
 DR InterPro: IPR001788; RNA\_dep\_RNAPol2.  
 DR InterPro: IPR000606; Viral\_helicase1.  
 DR Pfam: PF01443; Viral\_helicase1; 1.  
 DR SEQUENCE 1365 AA; 155129 MW; 35449D5E8026633 CRC64;  
 SQ

Query Match 73.1%; Score 19; DB 12; Length 1365;  
 Best Local Similarity 40.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXLD 10  
 DB 1302 KKTATAYALD 1311

RESULT 6  
 ID 08U248 PRELIMINARY; PRT; 1365 AA.  
 AC 08U248;  
 DT 01-MAR-2002 (TReMBLrel. 20, Created)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)  
 DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Putative, putative, putative, putative, and putative genes.  
 OS Bamboo mosaic virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potexvirus.  
 OX NCBI\_TaxID=35286;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BAMV-V;  
 RA Yang C.-C., Liu J.-S., Lin C.-P., Lin N.-S.;  
 RT "Nucleotide sequence and phylogenetic analysis of a bamboo mosaic  
 RT potexvirus isolate from common bamboo (Bambusa vulgaris McClure).";  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: L77962; AAL4030.1; -;  
 DR InterPro: IPR001788; RNA\_dep\_RNAPol2.  
 DR InterPro: IPR000606; Viral\_helicase1.  
 DR Pfam: PF00978; RNA\_dep\_RNAPol2; 1.  
 DR Pfam: PF01443; Viral\_helicase1; 1.  
 DR SEQUENCE 1365 AA; 155222 MW; B542C00E2ECFC990 CRC64;  
 SQ

Query Match 73.1%; Score 19; DB 12; Length 1365;  
 Best Local Similarity 40.0%; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXLD 10  
 DB 1302 KKTATAYALD 1311

RESULT 7  
 ID 09ZG37 PRELIMINARY; PRT; 46 AA.  
 AC 09ZG37;  
 DT 01-MAY-1999 (TReMBLrel. 10, Created)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
 DE Hypothetical 5.2 kDa protein (Fragment).  
 OS Chlamydia trachomatis.  
 OG Plasmid pCHL1.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=813;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN-L2 434B;  
 RA Wang L., Steenburg S.D., Zheng Y., Larsen S.H.;  
 RT "Gene identification of Chlamydia trachomatis by random DNA  
 RT sequencing.";  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF087330; AAD04105.1; -;  
 DR InterPro: IPR005350; UPF0137.  
 DR Pfam: PF03677; UPF0137; 1.  
 KW Hypothetical protein; Plasmid.  
 FT NON\_TER 1 46  
 FT NON\_TER 1 46  
 SQ SEQUENCE 46 AA; 5214 MW; D2FDAC41C1960867 CRC64;  
 SQ

Query Match 69.2%; Score 18; DB 2; Length 46;  
 Best Local Similarity 40.0%; Pred. No. 12;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXLD 10  
 DB 21 KKNQTAAASLD 30

RESULT 8  
 ID 08RTC9 PRELIMINARY; PRT; 108 AA.  
 AC 08RTC9;  
 DT 01-JUN-2002 (TReMBLrel. 21, Created)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
 DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Putative portal vertex protein GPQ (Fragment).  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CHO14;  
 RA Pradel N., Leroy-Setrin S., Lévrel V.;  
 RT "Sequences from Shiga toxin-producing Escherichia coli O91:H21.";  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF467512; AAL78348.1; -;  
 FT NON\_TER 1 108  
 FT NON\_TER 1 108  
 SQ SEQUENCE 108 AA; 12028 MW; 350B505A7A03B5B5 CRC64;  
 SQ

Query Match 69.2%; Score 18; DB 2; Length 108;  
 Best Local Similarity 40.0%; Pred. No. 25;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXLD 10  
 DB 54 KKSADLDD 63

RESULT 9  
 ID 09ANS7 PRELIMINARY; PRT; 153 AA.  
 AC 09ANS7;  
 DT 01-JUN-2001 (TReMBLrel. 17, Created)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)  
 DE 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE LuxT.  
 OS Vibrio harveyi.  
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.  
 OX NCBI\_TaxID=669;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lin Y.H., Miyamoto C.M., Melghe E.A.;  
 RT "Cloning and functional studies of a luxO regulator LuxT from Vibrio  
 RT harveyi.";  
 RL Biochim. Biophys. Acta 0:0-0(2001).  
 DR EMBL: AF321123; AAK09362.1; -;

SQ SEQUENCE 153 AA; 17475 MW; 33837BBA2A1819B CRC64;

Query Match 69.2%; Score 18; DB 2; Length 153;

Best Local Similarity 40.0%; Pred. No. 34;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXXLD 10

DB 54 KKTDFETALD 63

RESULT 10

O8MUVS PRELIMINARY; PRT; 160 AA.

AC O8MUVS; 01-MAR-2002 (TREMBLrel. 20, Created)

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=SKIN;

RA Strausberg R.;

RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: BC019336; AAL19336.1;

KW Hypothetical protein.

FT NON\_TER

SQ SEQUENCE 160 AA; 17041 MW; 9F7B497AAC06EDB8 CRC64;

Query Match 69.2%; Score 18; DB 4; Length 160;

Best Local Similarity 40.0%; Pred. No. 35;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXXLD 10

DB 52 KKSAAESMID 61

RESULT 11

O9KL32 PRELIMINARY; PRT; 161 AA.

AC O9KL32; 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

OS Transcriptional regulator, TetR family.

OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.

OX NCBI\_TaxID=666;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=EL TOR N16961 / SEROTYPE O1;

RA MEDLINE-20406833; PubMed-10952301;

RA Hedberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,

RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,

RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragol I., Sellers P.,

RA McDonald L., Ullrich T., Fleischmann R.D., Nielsen W.C., White O.,

RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,

RA Fraser C.M.;

RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio

RT cholerae."

RL Nature 406:477-483(2000).

DR EMBL: AE004419; AAF96814.1;

DR TIGR: AC0917;

KW Complete proteome.

SQ SEQUENCE 161 AA; 18749 MW; AA78003D18984288 CRC64;

Query Match 69.2%; Score 18; DB 16; Length 161;

Best Local Similarity 40.0%; Pred. No. 35;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXXLD 10

DB 62 KKTDFASALD 71

RESULT 12

O04541 PRELIMINARY; PRT; 171 AA.

AC O04541; 01-JUL-1997 (TREMBLrel. 04, Created)

DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)

DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC Euphorbia; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI\_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RA Osborn B.I., Vysotskaia V.S., Toriumi M., Yu G., Oji O., Buehler E.,

RA Conway A.B., Conway A.R., Dewar K., Feng J., Kim C., Kurtz D., Li Y.,

RA Shin F., Sun H., Davis R.W., Ecker J.R., Federspiel N.A.,

RA Theologis A.;

RT "The sequence of BAC F20P5 from Arabidopsis thaliana chromosome 1,"

RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RA Theologis A.;

RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL: AC002062; AAB61109.1;

DR InterPro: IPR003822; PAH.

DR Pfam: PF02671; PAH; 2.

SQ SEQUENCE 171 AA; 19992 MW; 691285E59734CA73 CRC64;

Query Match 69.2%; Score 18; DB 10; Length 171;

Best Local Similarity 40.0%; Pred. No. 37;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 KKKXXXXXXLD 10

DB 16 KKLTTTMDLD 25

RESULT 13

O9D8S4 PRELIMINARY; PRT; 205 AA.

AC O9D8S4; 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DE 1810038D15R1k protein (RIKEN CDNA 1810038D15 gene).

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=PANCREAS;

RA MEDLINE-21085660; PubMed-11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamazaki I.,

RA Saito T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,



RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Krenl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamilya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli U., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wyszynski B., Yoshida K., Hasegawa Y., Kawaji H., Kohzuki S.,  
 RA Hayashizaki Y.,  
 RT Functional annotation of a full-length mouse cDNA collection.\*;  
 RL Nature 409:685-690(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK007731; BAB25219.1; -  
 DR EMBL: BC003445; AAH03445.1; -  
 DR MGD: MGI:1916430; 1810038D15Rik.  
 DR InterPro: IPR000520; Exonuclease.  
 DR Pfam: PF00929; Exonuclease; 1.  
 DR SMART: SM00479; EXOIII; 1.  
 SQ SEQUENCE 205 AA; 23613 MW; D3FF9C701AB730E CRC64;  
 Query Match 69.2%; Score 18; DB 11; Length 205;  
 Best Local Similarity 40.0%; Pred. No. 43;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Oy 1 KKKXXXXXXD 10  
 Db 157 KKAASHRALD 166  
 RESULT 14  
 ID 08X427 PRELIMINARY; PRT; 212 AA.  
 AC 08X427;  
 DT 01-MAR-2002 (TReMBLrel. 20, Created)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Putative tet operon regulator.  
 GN YCDC OR Z1512 OR ECS1259.  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia  
 NCBI\_TaxID=83334;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE-21074935; PubMed-11206551;  
 RA Berra N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.V., Maynew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grodeck E.J., Davis N.W., Lim A., Dimailanta E.T., Potamousis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blatter F.R.;  
 RT "genome sequence of enterohaemorrhagic Escherichia coli O157:H7";  
 RL Nature 409:529-533(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-O157:H7 / RIMD 0509952;  
 RX MEDLINE-21156231; PubMed-11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
 RA Kuhara S., Shida T., Hattori M., Shimagawa H.;  
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli  
 O157:H7 and genomic comparison with a laboratory strain K-12";  
 RL DNA Res. 8:11-22(2001).  
 DR EMBL: AE005301; AAG55629.1; -.

DR EMBL: AP002554; BAB34682.1; -  
 DR InterPro: IPR001647; HTH\_Tetr.  
 DR Pfam: PF00440; tetr; 1.  
 DR PRINTS: PR00455; RHTEFTR.  
 KW Complete proteome.  
 SQ SEQUENCE 212 AA; 23703 MW; 5DA47B85C960F06C CRC64;  
 Query Match 69.2%; Score 18; DB 16; Length 212;  
 Best Local Similarity 40.0%; Pred. No. 45;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Oy 1 KKKXXXXXXD 10  
 Db 20 KKAILSALD 29  
 RESULT 15  
 ID 096609 PRELIMINARY; PRT; 215 AA.  
 AC 096609;  
 DT 01-MAY-1999 (TReMBLrel. 10, Created)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
 DE Surface antigen ariel1.  
 OS Entamoeba histolytica.  
 OC Eukaryota; Entamoebidae; Entamoeba.  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-HM1-TMSS;  
 RX MEDLINE-98084495; PubMed-9423879;  
 RA Mai Z., Samuelson J.;  
 RT "A new gene family (ariel) encodes asparagine-rich Entamoeba  
 RT histolytica antigens, which resemble the amebic vaccine candidate  
 RT serine-rich E. histolytica protein.";  
 RL Infect. Immun. 66:353-355(1998).  
 DR EMBL: AF093575; AAC72364.1; -  
 SQ SEQUENCE 215 AA; 23641 MW; 00B59B87816A5ECD CRC64;  
 Query Match 69.2%; Score 18; DB 5; Length 215;  
 Best Local Similarity 40.0%; Pred. No. 45;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Oy 1 KKKXXXXXXD 10  
 Db 50 KKSNSSELD 59  
 Search completed: June 5, 2003, 15:55:43  
 Job time : 32.0909 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:47:43 ; Search time 33.9394 Seconds  
(without alignments)  
39.261 Million cell updates/sec

Title: US-09-150-947F-15

Perfect score: 15

Sequence: 1 xxxxxxxxxd 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq\_101002:\*

1: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:\*

2: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:\*

3: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:\*

4: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:\*

5: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:\*

6: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:\*

7: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:\*

8: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:\*

9: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:\*

10: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:\*

11: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:\*

12: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:\*

13: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:\*

14: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:\*

15: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:\*

16: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:\*

17: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:\*

18: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:\*

19: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:\*

20: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:\*

21: /SID2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:\*

22: /SID2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:\*

23: /SID2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	40.0	2	2	AA10185
2	6	40.0	2	21	AA37936
3	6	40.0	3	5	AA40819
4	6	40.0	3	11	AA04607
5	6	40.0	3	13	AA00718
6	6	40.0	3	13	AA25615
7	6	40.0	3	14	AA36707
8	6	40.0	3	14	AA30728
9	6	40.0	3	14	AA30753
10	6	40.0	3	14	AA53144

11	6	40.0	3	15	AA44666
12	6	40.0	3	15	AA48960
13	6	40.0	3	16	AA82907
14	6	40.0	3	16	AA70472
15	6	40.0	3	16	AA64556
16	6	40.0	3	16	AA64722
17	6	40.0	3	17	AA11094
18	6	40.0	3	17	AA99823
19	6	40.0	3	17	AA99825
20	6	40.0	3	17	AA99827
21	6	40.0	3	18	AA39130
22	6	40.0	3	18	AA31143
23	6	40.0	3	18	AA25187
24	6	40.0	3	18	AA25193
25	6	40.0	3	18	AA25173
26	6	40.0	3	19	AA64738
27	6	40.0	3	19	AA71120
28	6	40.0	3	19	AA48589
29	6	40.0	3	19	AA56245
30	6	40.0	3	19	AA56225
31	6	40.0	3	19	AA56186
32	6	40.0	3	19	AA56197
33	6	40.0	3	20	AA50604
34	6	40.0	3	20	AA50605
35	6	40.0	3	20	AA43493
36	6	40.0	3	20	AA43494
37	6	40.0	3	20	AA39889
38	6	40.0	3	20	AA30549
39	6	40.0	3	20	AA30593
40	6	40.0	3	20	AA32860
41	6	40.0	3	20	AA32872
42	6	40.0	3	20	AA13741
43	6	40.0	3	20	AA88037
44	6	40.0	3	21	AA37920
45	6	40.0	3	21	AA37922

#### ALIGNMENTS

RESULT 1	AA10185
ID	AA10185 standard; Protein; 2 AA.
XX	
AC	AA10185;
XX	
DT	16-ARG-2002 (updated)
DT	14-ARG-1992 (first entry)
XX	
DE	Sequence of synthetic polypeptide for the prodn. of aspartame.
XX	
KW	Aspartyl-phenylalanine methyl ester; artificial sweetener;
XX	
OS	Synthetic.
XX	
PN	EP36258-A.
XX	
PD	23-SEP-1981.
XX	
PF	02-MAR-1981; 81EP-0300857.
XX	
PR	14-MAR-1980; 80US-0130462.
XX	
PA	(CETU-) CETUS CORP.
XX	
PI	Rose JE, White TV, Bahl CP;
XX	
DR	WPI; 1981-72133D/40.
XX	
DR	N-PDSB; AAN10059.
XX	
PT	Aspartame prodn. from synthetic polypeptide - produced by cloned microorganism

Platelet aggregati  
NL-3 truncated GA  
Non-RBD, non-YISGR  
Cancer metastasis  
RF-1 peptide 1 fro  
HPR3 peptide deriv  
Platelet-targetin  
Active domain from  
Mutant chemokine a  
Human IL-16 5' end  
Platelet-targeting  
LDV-peptide capabl  
IDA-peptide capabl  
RCD-peptide capabl  
Angiotensin II pep  
Peptide AII(1-3) u  
Integrin receptor  
Anti-inflammatory  
Anti-inflammatory  
Anti-inflammatory  
Anti-inflammatory  
Resin bound cyclic  
Resin bound cyclic  
Linker for dual av  
Linker for dual av  
Asparaginyl endope  
Amino acid sequenc  
Fibronectin protei  
Fibronectin protei  
RCD peptide. Unid  
Peptide used in th  
Trypsin modulating  
Trypsin modulating

XX PS Disclosure; Page 5; 18pp; English.  
CC CC The inventors claim a method for the prodn. of aspartame by first  
CC synthesizing double-stranded DNA in which a coding strand has  
CC alternating codons for Asp and the phe in sufficient number to  
CC produce a polypeptide which is stable in predetermined host  
CC microorganisms. The DNA strand is then inserted into a cloning  
CC vehicle so that the resulting chimera directs the synthesis of the  
CC Asp-phe protein. This protein is then cleaved with CNBr or trypsin  
CC by digestion with chymotrypsin.  
CC (Updated on 16-AUG-2002 to add missing OS field.)  
CC XX  
SQ Sequence 2 AA;  
Query Match 40.0%; Score 6; DB 2; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.4e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 10 D 10  
DB 1 D 1  
RESULT 2  
AAB37936  
ID AAB37936 standard; peptide; 2 AA.  
XX AC AAB37936;  
XX DT 27-FEB-2001 (first entry)  
XX DE Trypsin modulating oostatic factor (TMOF) peptide SEQ ID 41.  
XX KM Trypsin modulating oostatic factor; TMOF, pesticide; pest control;  
XX KW digestive enzyme synthesis inhibition; mosquito larvae; coleoptera;  
XX KW lepidopteran; dipteran; blood-sucking insect; Diptera; Nematocera;  
XX KW Colicidae; Culicidae; Corethrinae; Ceratopogonidae; Simuliidae.  
OS Synthetic.  
XX XX  
XX WO200063233-A2.  
XX PD 26-OCT-2000.  
XX PF 18-APR-2000; 2000WO-US10235.  
XX PR 21-APR-1999; 99US-0295996.  
XX PA (UVFL ) UNIV FLORIDA RES FOUND INC.  
XX PA (UYNC-) UNIV NORTH CAROLINA STATE.  
XX PI Borovsky D, Linderman RJ;  
XX DR WPI: 2000-687156/67.  
XX PT Novel trypsin modulating oostatic factor compound peptides, useful for  
XX PT pest control including pests of agricultural crops -  
XX PS Claim 3; Page 33; 50pp; English.  
XX This invention relates to a trypsin modulating oostatic factor (TMOF)  
CC compound. The compound specifically does not include the TMOF analogue  
CC peptides represented by AAB37942 - AAB37949. Sequences AAB37900 -  
CC AAB37941 are TMOF analogues which may be included in the composition. The  
CC TMOF compound has pesticidal activity and is an inhibitor of pest  
CC digestive enzyme synthesis. The TMOF compound is useful for controlling  
CC pests such as mosquito larvae, coleopterans, lepidopterans, dipterans or  
CC blood-sucking insects of order Diptera, suborder Nematocera, family  
CC Colicidae or subfamily Culicidae, Corethrinae, Ceratopogonidae and  
CC Simuliidae.

SQ Sequence 2 AA;  
Query Match 40.0%; Score 6; DB 21; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.4e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 10 D 10  
DB 2 D 2  
RESULT 3  
AAP40819  
ID AAP40819 standard; peptide; 3 AA.  
XX AC AAP40819;  
XX DT 03-AUG-1992 (first entry)  
XX DE Sequence of fragment D corresp. to residues 25-27 of human  
XX DE pancreatic growth hormone releasing factor (somatotrinin) (hGRF).  
XX DE Hormone; dwarfism; therapy; retarded growth;  
XX KW anabolic protein deficiency; growth promoter; lactation.  
XX OS Homo sapiens.  
XX PH  
XX PH Key Location/Qualifiers  
FT Modified-site 1 /label= Boc-D(OBz1)  
FT /note= "Boc-tertlobutylloxycarbonyl (carbamate);  
FT OBz1-ester benzyllyque"  
FT Modified-site 3 /label= M-NH-NH2  
XX XX  
XX AU8424774-A.  
XX PD 30-AUG-1984.  
XX PF 20-FEB-1984; 84AU-0400343.  
XX PR 29-NOV-1983; 83FR-0019058.  
XX PR 21-FEB-1983; 83FR-0002781.  
XX PA (SNFI ) SANOFI SA.  
XX PI Diaz J, Demarne H, Roncucci R, Schmelck PH;  
XX DR WPI: 1984-256760/42.  
XX PT Synthesis of hGRF in liquid phase reactions - with use of new  
XX PT peptide fragments  
XX PS Claim 4; Page 67; 76pp; French.  
XX XX  
XX The inventors claim hGRF fragments used for the synthesis of hGRF.  
CC using the method somatotrinin can be obt. on a large scale with  
CC good yield and good purity. The prod. is used in man for the  
CC treatment of dwarfism and retarded growth and for anabolic protein  
CC deficiencies. In animals it is useful for promoting wt. growth for  
CC increased prodn., lactation etc. This index is based on Ep-122818,  
CC which is the equivalent of AU8424774.  
XX SQ  
SQ Sequence 3 AA;  
Query Match 40.0%; Score 6; DB 5; Length 3;  
Best Local Similarity 100.0%; Pred. No. 7.4e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 10 D 10  
DB 1 D 1

```

RESULT 4
AAR04607
ID AAR04607 standard; peptide; 3 AA.
XX
AC AAR04607;
XX
DE 05-SEP-1990 (first entry)
XX
DE Antiviral agent.
XX
KM Antiviral; M2; poliovirus; polio; hepatitis.
XX
OS Synthetic.
XX
PN JP02078631-A.
XX
PD 19-MAR-1990.
XX
PF 14-SEP-1988; 88JP-0228843.
XX
PR 14-SEP-1988; 88JP-0228843.
XX
PA (NIRA ) NIPPON MINING KK.
XX
DR WPI; 1990-129060/17.
XX
PT Antiviral agent contg. tripeptide (unit) -
PT of basic aminoacid, then alanine, glycine or sarcosine, and
PT acidic aminoacid, effective against virus with protein-terminated DNA
PT or RNA.
XX
PS Disclosure; 4pp; Japanese.
XX
CC Peptide is effective against inhibiting propagation of DNA or RNA
CC bonded, protein containing viruses eg. Poliovirus, Hepatitis virus.
XX
SQ Sequence 3 AA;

Query Match 40.0%; Score 6; DB 11; Length 3;
Best Local Similarity 100.0%; Pred. No. 7.4e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10
DB 3 D 3

RESULT 5
AAR00718
ID AAR00718 standard; peptide; 3 AA.
XX
AC AAR00718;
XX
DE 29-MAY-1990 (first entry)
XX
DE Core repeat of cell-adhesive protein.
XX
KM Cell adhesion; anti-metastatic agent; immunomodulation; core repeat.
XX
PN EP347931-A.
XX
PD 27-DEC-1989.
XX
PF 23-JUN-1989; 89EP-0111468.
XX
PR 24-JUN-1988; 88JP-0156133.
XX
PA (AZUMA/) AZUMA I.
XX
PI Saito I, Nishi N, Azuma I, Tokura S;
XX
DR WPI; 1990-001305/01.

```

```

XX
XX Polypeptide with repeated sequences of cell adhesion protein used
PT as anti-metastatic agent for cancer and agonist or antagonist of
PT cell-adhesion proteins
XX
PS Claim 2; page 14; 16pp; English.
XX
CC Peptide core is repeated 2-20 times to form a cell-adhesive protein of
CC mol. wt. 1,500-5,000. The protein is an (ant)agonist of cell-adhesive
CC proteins such as fibronectin. It has high antimetastatic activity
CC against cancer and can be used in immunomodulation, wound healing,
CC platelet aggregation inhibition and alleviation of neuro-disorders.
CC See also AAR00722.
XX
SQ Sequence 3 AA;

Query Match 40.0%; Score 6; DB 11; Length 3;
Best Local Similarity 100.0%; Pred. No. 7.4e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10
DB 3 D 3

RESULT 6
AAR25615
ID AAR25615 standard; peptide; 3 AA.
XX
AC AAR25615;
XX
DE 20-JAN-1993 (first entry)
XX
DE Harvey Ras inhibitor - Ras 30-32.
XX
KM GTPase; activating protein; GAP; oncogenesis; cancer; truncated;
KM colorectal; exocrine pancreatic; myeloid leukaemia; adenyl acylase.
XX
OS Synthetic.
XX
PN EP496162-A.
XX
PD 29-JUL-1992.
XX
PF 16-DEC-1991; 91EP-0311658.
XX
PR 24-DEC-1990; 90US-0632891.
XX
PA (MERI ) MERCK & CO INC.
XX
PI Freidinger RM, Garsky VM, Gibbs JB, Schnaber MD, Sigal IS;
XX
DR WPI; 1992-251427/31.
XX
PT Synthetic peptide inhibitors of Ras-GAP interaction - as
PT antitumour agents against e.g. colorectal carcinoma, exocrine
PT pancreatic carcinoma and myeloid leukaemia
XX
PS Claim 3; page 7; 31pp; English.
XX
CC The peptide corresponds to a C- and N-terminal truncated version of
CC the prod. of the Harvey ras oncogene from residues 17-44, i.e. from
CC residues 30-32. The peptide is capable of binding to GTPase
CC activating protein (GAP) thus inhibiting Ras activation by GAP and
CC inhibiting Ras-dependent cell transformation. Derivs. of the Ras
CC peptide are also provided, comprising C-terminal truncations, N-
CC terminal truncations, or modifications of specific amino acids, e.g.
CC the most potent Ras inhibitor is Ras 17-32. Oncogenic forms of Ras
CC occur in various human cancers, e.g. colorectal carcinoma, exocrine
CC pancreatic carcinoma and myeloid leukaemias. Ras complexed with GTP
CC can bind GAP and this is thought to be the important step in the
CC transformation activity of Ras. The peptide is also capable of
CC inhibiting GAP as well as the interaction of Ras with yeast

```

CC adenylyl acylase. See also AAR25602-29.  
 XX  
 SQ Sequence 3 AA;

Query Match 40.0%; Score 6; DB 13; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
 |  
 Db 1 D 1

RESULT 7  
 AAR36707  
 ID AAR36707 standard; peptide: 3 AA.  
 XX

AC AAR36707;

XX 26-AUG-1993 (first entry)

DE Adhesion formation prevention RGD-contg. peptide.

XX Tissue repair; peritoneum; surgery; post-surgically; inhibition;  
 KW platelet aggregation; cardiovascular; orthopedic; thoracic;

KW ophthalmic; CNS; use.

XX Synthetic.

XX WO9308818-A.

XX 13-MAY-1993.

XX 06-NOV-1992; 92WO-US09494.

XX 07-NOV-1991; 91US-0789231.

XX (UYSC-) UNIV SOUTHERN CALIFORNIA.

XX Dizerega GS, Rodgers KE;

XX WPI; 1993-167381/20.

XX Prevention of adhesion formation, partic. post-surgically - comprises  
 PT administering a RGD-contg. peptide for a time sufficient to permit  
 PT tissue repair

PS Example; Page 18; 22pp; English.

XX The sequence is that of an RGD-contg. peptide which is used in a  
 CC method for prevention of adhesion formation for a time sufficient  
 CC to permit tissue repair. The method is used for minimizing or  
 CC preventing adhesion formation, partic. in the peritoneum following  
 CC surgery, but also for e.g. cardiovascular, orthopedic, thoracic,  
 CC ophthalmic, CNS and other uses. In addn., the peptide inhibits  
 CC platelet aggregation and does not induce inflammation or trauma  
 CC at the site of administration.

XX Sequence 3 AA;

Query Match 40.0%; Score 6; DB 14; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
 |  
 Db 3 D 3

RESULT 8  
 AAR30728  
 ID AAR30728 standard; peptide: 3 AA.  
 XX

AC AAR30728;  
 XX  
 DT 20-MAY-1993 (first entry)  
 XX

DE IGE-mast cell reaction preventing peptide intermediate.

XX Basophil; protective group; fragment condensing.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-t-butoxycarbonyl-beta-benzyl-L-aspartyl"

XX JP04360898-A.

XX 14-DEC-1992.

XX 04-JUN-1991; 91JP-0159394.

XX 04-JUN-1991; 91JP-0159394.

XX (TANA) TANABE SEIYAKU CO.

XX WPI; 1993-032710/04.

XX L-Aspartyl-L-seryl-L-aspartyl-L-prolyl-L-arginine prepn. -

PT involves fragment-condensing tri-peptide with amino-protected

PT di-peptide to form penta-peptide

XX Claim 7; Page 2; 8pp; Japanese.

XX The prepn. of DSDPR is claimed and involves e.g. fragment-condensing

CC C-protected tripeptide with N-protected dipeptide to form pentapeptide

CC and then removing the protective gps.

CC N-t-butoxycarbonyl-beta-benzyl-L-aspartyl-L-prolyl-L-arginine is a

CC new cpd."

XX Sequence 3 AA;

QY 10 D 10  
 |  
 Db 1 D 1

RESULT 9  
 AAR30753  
 ID AAR30753 standard; peptide: 3 AA.  
 XX

AC AAR30753;

XX 26-MAY-1993 (first entry)

XX IGE-mast cell reaction inhibitory peptide prepn. peptide.

XX Physiologically active; basophil; tripeptide.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-t-butoxycarbonyl-beta-benzyl-L-Asp"

FT Modified-site 5 /note= "N(G)-nitro-L-Arg benzyl ester"

XX JP04360899-A.

XX 14-DEC-1992.

PF 04-JUN-1991; 91JP-0159492.  
 XX  
 PR 04-JUN-1991; 91JP-0159492.  
 XX  
 PA (TANA) TANABE SEIYAKU CO.  
 XX  
 DR WPI; 1993-032711/04.  
 XX  
 PT Prepn. of physiologically active penta-peptide - comprises fragment-  
 PT condensing specified carboxyl protected tri-peptide and amino-  
 PT protected di-peptide and removing protective gps.  
 XX  
 PS Claim; Page 2; 8pp; Japanese.  
 XX  
 CC The peptide is used in the prepn. (claimed) of a physiologically.  
 CC active and prevents the Ige - mast cell (basophil) reaction.  
 XX  
 SQ Sequence 3 AA;

Query Match 40.0%; Score 6; DB 14; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
 DB 1 D 1

RESULT 10  
 AAR53144  
 ID AAR53144 standard; peptide; 3 AA.  
 XX  
 AC AAR53144;  
 XX  
 DT 02-JUN-1994 (first entry)  
 XX  
 DE RGD peptide derivative #6.  
 XX  
 KM Drug; organ transplantation; rejection; immune disorder;  
 KM systemic lupus.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /note= "C9H19CO-Arg, C13H27CO-Arg, C15H31CO-Arg,  
 FT CH3-[(CH(CH3))- (CH2)3]3-CH(CH3)-CH2-CO-Arg or  
 FT CH3-[(CH(CH3))- (CH2)3]2-CH(CH3)-CH2-CO-Arg"  
 FT Modified-site 4  
 FT /note= "Asp-OH or Asp-NH2"  
 FT  
 FT  
 FT  
 PN JP05255105-A.  
 XX  
 PD 05-OCT-1993.  
 XX  
 PF 16-MAR-1992; 92JP-0058460.  
 XX  
 PR 16-MAR-1992; 92JP-0058460.  
 XX  
 PA (FUJI) FUJI PHOTO FILM CO LTD.  
 XX  
 DR WPI; 1993-348360/44.  
 XX  
 PT Immuno-control drug for organ transplant rejection etc. - contains  
 PT peptide having arginine, glycine, aspartic acid sequence  
 XX  
 PS Disclosure; Page 3; 11pp; Japanese.  
 XX  
 CC The sequences given in AAR44043-47 and AAR53144 represent examples of  
 CC the claimed RGD containing peptide of the invention. These peptides all  
 CC correspond to the generic formulae:  
 CC HO2-(CH2)m-C(O)-(X)-Arg-Gly-Asp-(Y)n-O-CH2CH(OR1)CH2OR2 or  
 CC R3-[(X)-Arg-Gly-Asp-(Y)n-Z

CC [X], [Y] = amino acid or peptide residues;  
 CC m = 1-5;  
 CC n = 1-5;  
 CC R1, R2 = H or 8-24C acyl or alkyl;  
 CC R3 = 6-24C acyl;  
 CC Z = hydroxyl or amino.  
 CC These peptides form the active part of drugs which are used for the  
 CC control of organ transplantation rejection or immune disorders such  
 CC as systemic lupus.  
 XX  
 SQ Sequence 3 AA;

Query Match 40.0%; Score 6; DB 14; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
 DB 3 D 3

RESULT 11  
 AAR44666  
 ID AAR44666 standard; peptide; 3 AA.  
 XX  
 AC AAR44666;  
 XX  
 DT 20-JAN-1995 (first entry)  
 XX  
 DE Platelet aggregation or adhesion inhibitor - peptide 5.  
 XX  
 KM Platelet aggregation; adhesion; inhibitor; guest; host;  
 KM beta-cyclodextrin; protease; resistance; degradation.  
 XX  
 OS Synthetic.  
 XX  
 PN JP06116289-A.  
 XX  
 PD 26-APR-1994.  
 XX  
 PF 09-OCT-1992; 92JP-0271294.  
 XX  
 PR 09-OCT-1992; 92JP-0271294.  
 XX  
 PA (FUJI) FUJI PHOTO FILM CO LTD.  
 XX  
 DR WPI; 1994-173759/21.  
 XX  
 PT Complex of adhesion peptide in host molecule e.g.  
 PT beta-cyclodextrin - useful as platelet aggregation inhibitor  
 PT which is resistant to protease degradation in-vivo  
 XX  
 PS Disclosure; Page 4; 5pp; Japanese.  
 XX  
 CC A peptide complex contg. peptide 5 as guest mol. is useful as a  
 CC platelet aggregation or adhesion inhibitor. It is hardly hydrolysed  
 CC by protease in vivo and thus maintains its effect for a long period.  
 XX  
 SQ Sequence 3 AA;

Query Match 40.0%; Score 6; DB 15; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
 DB 3 D 3

RESULT 12  
 AAR48960  
 ID AAR48960 standard; Protein; 3 AA.  
 XX

AC AAR48960;  
 XX  
 DT 12-SEP-1994 (first entry)  
 XX  
 DE NL4-3 truncated GAG C-terminal peptide.  
 XX  
 KM HIV-1; HXB2; antisense; sequence inversion; antisense virus; infection;  
 KM naturally occurring virus; NOV; translation; replication; infectivity;  
 KM hepatitis B; HIV-2; SIV; flip-over PCR.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9403596-A.  
 XX  
 PD 17-FEB-1994.  
 XX  
 PF 30-JUL-1993; 93WO-US07179.  
 XX  
 PR 30-JUL-1992; 92US-0921104.  
 XX  
 PA (UYHA-) UNIV HAWAII.  
 XX  
 PI Hu W, Wang J;  
 XX  
 DR WPI: 1994-065685/08.  
 XX  
 DR N-PSDB: AA057688.  
 XX  
 PT New antisense viruses and anti-sense-ribozyme viruses - used for  
 PT treating or preventing viral infections, partic. HIV-1, HIV-2 or  
 PT SIV infection  
 XX  
 PS Disclosure: Page 108; 167pp: English.  
 XX  
 CC This sequence is encoded by a PCR fragment of NL4-3 and represents the  
 CC C-terminal peptide fragment of the truncated GAG protein. The DNA  
 CC encoding this fragment was ligated into ClaI/SalI digested pX and the  
 CC corresponding plasmid was used to produce the antisense virus of the  
 CC invention. Antisense or truncated RNAs expressed by these viruses  
 CC bind to the mRNAs expressed by the naturally occurring viruses (NOVs)  
 CC and prevent the mRNAs from being translated into proteins, thereby  
 CC preventing the NOV from replicating. The antisense viruses maintain  
 CC the infectivity of the NOVs, allowing antisense RNAs to reach the  
 CC mRNAs of the natural viruses. Antisense viruses such as these may be  
 CC used for treating or preventing a viral infection, particularly HIV-1,  
 CC HIV-2 or SIV infection or hepatitis B infection.  
 CC  
 SQ Sequence 3 AA;  
 Query Match 40.0%; Score 6; DB 15; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 10 D 10  
 |  
 Db 2 D 2  
 2 D 2

RESULT 13  
 AAR82907  
 ID AAR82907 standard; peptide: 3 AA.  
 XX  
 AC AAR82907;  
 XX  
 DT 20-DEC-1995 (first entry)  
 XX  
 DE Non-RGD, non-YISGR cancer metastasis inhibitory peptide #1.  
 XX  
 KM Cancer metastasis; adhesive peptide; core sequence; dextran; cancer;  
 KM water soluble polysaccharide; metastasis; wound; immunogenicity.  
 XX  
 OS Synthetic.  
 XX  
 PN JP07089999-A.

XX  
 PD 04-APR-1995.  
 XX  
 PF 17-SEP-1993; 93JP-0254779.  
 XX  
 PR 17-SEP-1993; 93JP-0254779.  
 XX  
 PA (JABG ) NIPPON ZEON KK.  
 XX  
 DR WPI: 1995-167254/22.  
 XX  
 PT Cancer metastasis inhibitive peptide derivs. - useful for inhibition  
 PT of cancer metastasis, healing of wounds and regulation of  
 PT immunogenicity.  
 XX  
 PS Disclosure: Page 3; 6pp: Japanese.  
 XX  
 CC The peptides AAR70472-90 and AAR82907-24 are peptide derivatives which  
 CC inhibit cancer metastasis. They are composed of an adhesive peptide  
 CC with a core sequence selected from: RGD (AAR70472-85), YIGSR  
 CC (AAR70486-90) or other sequence (AAR82907-24), linked to a water soluble  
 CC polysaccharide, preferably a water soluble dextran, at the C-terminus.  
 CC The peptides are useful in inhibiting cancer metastasis, healing wounds  
 CC and the regulation of immunogenicity.  
 CC  
 SQ Sequence 3 AA;  
 Query Match 40.0%; Score 6; DB 16; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 10 D 10  
 |  
 Db 2 D 2  
 2 D 2

RESULT 14  
 AAR70472  
 ID AAR70472 standard; peptide: 3 AA.  
 XX  
 AC AAR70472;  
 XX  
 DT 20-DEC-1995 (first entry)  
 XX  
 DE Cancer metastasis inhibitory peptide core RGD sequence.  
 XX  
 KM Cancer metastasis; adhesive peptide; core sequence; dextran; cancer;  
 KM water soluble polysaccharide; metastasis; wound; immunogenicity.  
 XX  
 OS Synthetic.  
 XX  
 PN JP07089999-A.  
 XX  
 PD 04-APR-1995.  
 XX  
 PF 17-SEP-1993; 93JP-0254779.  
 XX  
 PR 17-SEP-1993; 93JP-0254779.  
 XX  
 PA (JABG ) NIPPON ZEON KK.  
 XX  
 DR WPI: 1995-167254/22.  
 XX  
 PT Cancer metastasis inhibitive peptide derivs. - useful for inhibition  
 PT of cancer metastasis, healing of wounds and regulation of  
 PT immunogenicity.  
 XX  
 PS Disclosure: Page 2; 6pp: Japanese.  
 XX  
 CC The peptides AAR70472-90 and AAR82907-24 are peptide derivatives which  
 CC inhibit cancer metastasis. They are composed of an adhesive peptide  
 CC with a core sequence selected from: RGD (AAR70472-85), YIGSR  
 CC (AAR70486-90) or other sequence (AAR82907-24), linked to a water soluble  
 CC polysaccharide, preferably a water soluble dextran, at the C-terminus.  
 CC The peptides are useful in inhibiting cancer metastasis, healing wounds  
 CC and the regulation of immunogenicity.



CC polysaccharide, preferably a water soluble dextran, at the C-terminus.  
 CC The peptides are useful in inhibiting cancer metastasis, healing wounds  
 CC and the regulation of immunogenicity.

XX Sequence 3 AA;

Query Match 40.0%; Score 6; DB 16; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10

DB 3 D 3

# RESULT 15

ID AAR64556 standard; Peptide; 3 AA.

AC AAR64556;

DT 01-SEP-1995 (first entry)

DE RF-1 peptide 1 from respiratory syncytial virus.

KW antiviral activity; DP-178; DP-107; diagnostic; HIV-1LAI;

KW human immunodeficiency virus; transmembrane protein; gp41;

KW alpha helix; leucine zipper; DP-185; respiratory syncytial virus;

OS RSV.

XX Synthetic.

XX Key

FT Modified-site

FT 1

FT Location/Qualifiers

FT /note= "optionally has an amino, acetyl,

FT 9-fluorenylmethoxy-carbonyl, hydrophobic or

FT macromolecular carrier gp. attached"

FT Modified-site

FT 3

FT /note= "optionally has a carboxyl, amide, hydrophobic

FT or macromolecular carrier gp. attached"

XX WO9428920-A.

XX 22-DEC-1994.

XX 07-JUN-1994; 94WO-US05739.

XX 07-JUN-1993; 93US-0073028.

XX (UYDU-) UNIV DUKE.

XX Barney SO, Bolognesi DP, Lambert DM, Matthews TJ;

XX Petteway SR, Wild CT;

XX WPI; 1995-036105/05.

XX Computer search generated synthetic peptides - are inhibitors of

XX HIV transmission

XX Claim 14; Page 137; 182pp; English.

XX Sequence 3 AA;

Query Match 40.0%; Score 6; DB 16; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10

DB 3 D 3

Search completed: June 5, 2003, 16:00:25  
 Job time : 35.9394 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:23:33 ; Search time 8.78788 Seconds  
(without alignments)  
33.481 Million cell updates/sec

Title: US-09-150-947F-15  
Perfect score: 15  
Sequence: 1 XXXXXXXXXXD 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCRTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6	40.0	2	1	US-07-820-154A-15 Sequence 15, Appl
2	6	40.0	2	1	US-07-820-154A-23 Sequence 23, Appl
3	6	40.0	2	1	US-07-820-154A-33 Sequence 33, Appl
4	6	40.0	2	1	US-07-820-154A-39 Sequence 39, Appl
5	6	40.0	2	1	US-08-358-160-171 Sequence 171, Appl
6	6	40.0	2	1	US-08-153-799-20 Sequence 20, Appl
7	6	40.0	2	2	US-08-097-554A-15 Sequence 15, Appl
8	6	40.0	2	2	US-08-097-554A-23 Sequence 23, Appl
9	6	40.0	2	2	US-08-097-554A-33 Sequence 33, Appl
10	6	40.0	2	2	US-08-097-554A-39 Sequence 39, Appl
11	6	40.0	2	2	US-08-539-433-15 Sequence 15, Appl
12	6	40.0	2	2	US-08-539-433-16 Sequence 16, Appl
13	6	40.0	2	2	US-08-539-433-17 Sequence 17, Appl
14	6	40.0	2	2	US-09-327-424-2 Sequence 2, Appl
15	6	40.0	2	3	US-08-480-640A-15 Sequence 15, Appl
16	6	40.0	2	3	US-08-480-640A-23 Sequence 23, Appl
17	6	40.0	2	3	US-08-480-640A-33 Sequence 33, Appl
18	6	40.0	2	3	US-08-480-640A-39 Sequence 39, Appl
19	6	40.0	2	3	US-08-295-802-15 Sequence 15, Appl
20	6	40.0	2	3	US-08-295-802-23 Sequence 23, Appl
21	6	40.0	2	3	US-08-295-802-33 Sequence 33, Appl
22	6	40.0	2	3	US-08-295-802-39 Sequence 39, Appl
23	6	40.0	2	4	US-09-326-335-2 Sequence 2, Appl
24	6	40.0	2	4	US-08-488-237A-15 Sequence 15, Appl
25	6	40.0	2	4	US-08-488-237A-23 Sequence 23, Appl
26	6	40.0	2	4	US-08-488-237A-33 Sequence 33, Appl
27	6	40.0	2	4	US-08-488-237A-39 Sequence 39, Appl

28	6	40.0	2	4	US-08-375-992A-15 Sequence 15, Appl
29	6	40.0	2	4	US-08-375-992A-23 Sequence 23, Appl
30	6	40.0	2	4	US-08-375-992A-33 Sequence 33, Appl
31	6	40.0	2	4	US-08-375-992A-39 Sequence 39, Appl
32	6	40.0	2	4	US-09-295-996B-23 Sequence 23, Appl
33	6	40.0	2	4	US-09-295-996B-35 Sequence 35, Appl
34	6	40.0	2	4	US-09-295-996B-58 Sequence 58, Appl
35	6	40.0	2	5	PCR-US93-00324-15 Sequence 15, Appl
36	6	40.0	2	5	PCR-US93-00324-23 Sequence 23, Appl
37	6	40.0	2	5	PCR-US93-00324-33 Sequence 33, Appl
38	6	40.0	2	5	PCR-US93-00324-39 Sequence 39, Appl
39	6	40.0	3	1	US-07-748-943-2 Sequence 2, Appl
40	6	40.0	3	1	US-07-780-790A-1 Sequence 1, Appl
41	6	40.0	3	1	US-08-169-524-6 Sequence 6, Appl
42	6	40.0	3	1	US-08-081-539-117 Sequence 117, Appl
43	6	40.0	3	1	US-08-251-027-1 Sequence 1, Appl
44	6	40.0	3	1	US-08-251-027-2 Sequence 2, Appl
45	6	40.0	3	1	US-08-251-027-3 Sequence 3, Appl

#### ALIGNMENTS

RESULT 1  
US-07-820-154A-15  
Sequence 15, Application US/07820154A  
Patent No. 5382425  
GENERAL INFORMATION:  
APPLICANT: Cochran Ph.D., Mark D  
APPLICANT: Junker M.S., David E  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/820,154A  
FILING DATE: 19920113  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)977-9550  
TELEFAX: (212)664-0525  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-820-154A-15  
Query Match 40.0%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 10 D 10  
Db 1 D 1  
RESULT 2  
US-07-820-154A-23

Sequence 23, Application US/07820154A  
Patent No. 5382425  
GENERAL INFORMATION:  
APPLICANT: Cochran Ph.D., Mark D  
APPLICANT: Junker M.S., David E  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/820.154A  
FILING DATE: 19920113  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)977-9550  
TELEFAX: (212)664-0525  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-820-154A-23

Query Match 40.0%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
DB 1 D 1

RESULT 3  
US-07-820-154A-33  
Sequence 33, Application US/07820154A  
Patent No. 5382425  
GENERAL INFORMATION:  
APPLICANT: Cochran Ph.D., Mark D  
APPLICANT: Junker M.S., David E  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/820.154A  
FILING DATE: 19920113  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)977-9550  
TELEFAX: (212)664-0525  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-820-154A-33

Query Match 40.0%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
DB 1 D 1

RESULT 4  
US-07-820-154A-39  
Sequence 39, Application US/07820154A  
Patent No. 5382425  
GENERAL INFORMATION:  
APPLICANT: Cochran Ph.D., Mark D  
APPLICANT: Junker M.S., David E  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/820.154A  
FILING DATE: 19920113  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)977-9550  
TELEFAX: (212)664-0525  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-820-154A-39

Query Match 40.0%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
DB 1 D 1

RESULT 5  
US-08-358-160-171  
Sequence 171, Application US/08358160  
Patent No. 5663143

GENERAL INFORMATION:  
APPLICANT: LEY, Arthur C.  
APPLICANT: LADNER, Robert C.  
APPLICANT: GUTERMAN, Sonia K.  
APPLICANT: ROBERTS, Bruce L.  
APPLICANT: MARKLAND, William  
APPLICANT: KENT, Rachel B.  
TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ  
NUMBER OF SEQUENCES: 234  
TITLE OF INVENTION: DOMAINS THAT INHIBIT HUMAN NEUTROPHIL ELASTASE  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROMDY AND NETMARK  
STREET: 419 Seventh Street, N.W. Suite 300  
CITY: Washington  
STATE: District of Columbia  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/358,160  
FILING DATE: 16-DEC-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/133,031  
FILING DATE: 13-OCT-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/009,219  
FILING DATE: 26-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/664,989  
FILING DATE: 01-MAR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/487,063  
FILING DATE: 02-MAR-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/240,160  
FILING DATE: 02-SEP-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Cooper, Iver P.  
REGISTRATION NUMBER: 28, 005  
REFERENCE/DOCKET NUMBER: LEY-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 171:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-358-160-171

Query Match 40.0%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
DB 1 D 1

RESULT 6  
US-08-153-799-20  
Sequence 20, Application US/08153799  
Patent No. 576883  
GENERAL INFORMATION:  
APPLICANT: Ballance, David J

APPLICANT: Goodey, Andrew R  
TITLE OF INVENTION: Polypeptides  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: R Hain Swope, BOC Health Care Inc  
STREET: 100 Mountain Avenue  
CITY: Murray Hill  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07974  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/153,799  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/847975  
FILING DATE: 06-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 8909916.2  
FILING DATE: 29-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB90/00650  
FILING DATE: 26-APR-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/775952  
FILING DATE: 29-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Swope, R Hain  
REGISTRATION NUMBER: 24864  
REFERENCE/DOCKET NUMBER: 92H832  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 665 2400  
TELEFAX: (908) 771 6159  
TELEX: 219484  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-153-799-20

Query Match 40.0%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
DB 1 D 1

RESULT 7  
US-08-097-554A-15  
Sequence 15, Application US/08097554A  
Patent No. 586312  
GENERAL INFORMATION:  
APPLICANT: Cochran Ph.D., Mark D  
APPLICANT: Junker M.S., David E  
TITLE OF INVENTION: Recombinant Symplocos Virus  
NUMBER OF SEQUENCES: 112  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,554A  
FILING DATE: July 22, 1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)977-9550  
TELEFAX: (212)664-0525  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-097-554A-15

Query Match 40.0%; Score 6; DB 2; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
DB 1 D 1

RESULT 8  
US-08-097-554A-23  
Sequence 23, Application US/08097554A  
Patent No. 5869312  
GENERAL INFORMATION:  
APPLICANT: Cochran Ph.D., Mark D  
APPLICANT: Junker M.S., David E  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 112  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,554A  
FILING DATE: July 22, 1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)977-9550  
TELEFAX: (212)664-0525  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-097-554A-23

Query Match 40.0%; Score 6; DB 2; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
DB 1 D 1

RESULT 9  
US-08-097-554A-33  
Sequence 33, Application US/08097554A  
Patent No. 5869312  
GENERAL INFORMATION:  
APPLICANT: Cochran Ph.D., Mark D  
APPLICANT: Junker M.S., David E  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 112  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,554A  
FILING DATE: July 22, 1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)977-9550  
TELEFAX: (212)664-0525  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-097-554A-33

Query Match 40.0%; Score 6; DB 2; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
DB 1 D 1

RESULT 10  
US-08-097-554A-39  
Sequence 39, Application US/08097554A  
Patent No. 5869312  
GENERAL INFORMATION:  
APPLICANT: Cochran Ph.D., Mark D  
APPLICANT: Junker M.S., David E  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 112  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,554A  
FILING DATE: July 22, 1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)977-9550  
TELEFAX: (212)664-0525  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-097-554A-39

Query Match 40.0%; Score 6; DB 2; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
DB 1 D 1

RESULT 11  
US-08-539-432-15  
Sequence 15, Application US/08539432  
Patent No. 5872210  
GENERAL INFORMATION:  
APPLICANT: MEDABALIMI, JOHN L.  
TITLE OF INVENTION: TRANSFRAME INHIBITORY  
TITLE OF INVENTION: ELEMENT OF VIRAL  
TITLE OF INVENTION: PROTEASE  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & PINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10054  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/539,432  
FILING DATE: 05-OCT-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4194  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2  
TYPE: Amino Acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Linear  
US-08-539-432-15

Query Match 40.0%; Score 6; DB 2; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
DB 1 D 1

RESULT 12  
US-08-539-432-16  
Sequence 16, Application US/08539432  
Patent No. 5872210  
GENERAL INFORMATION:  
APPLICANT: MEDABALIMI, JOHN L.  
TITLE OF INVENTION: TRANSFRAME INHIBITORY  
TITLE OF INVENTION: ELEMENT OF VIRAL  
TITLE OF INVENTION: PROTEASE  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & PINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10054  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/539,432  
FILING DATE: 05-OCT-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4194  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2  
TYPE: Amino Acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Linear  
US-08-539-432-16

Query Match 40.0%; Score 6; DB 2; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
DB 1 D 1

RESULT 13  
US-08-539-432-17  
Sequence 17, Application US/08539432  
Patent No. 5872210  
GENERAL INFORMATION:  
APPLICANT: MEDABALIMI, JOHN L.  
TITLE OF INVENTION: TRANSFRAME INHIBITORY  
TITLE OF INVENTION: ELEMENT OF VIRAL  
TITLE OF INVENTION: PROTEASE  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & PINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10054

Query Match 40.0%; Score 6; DB 2; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/539,432  
FILING DATE: 05-OCT-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4194  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2  
TYPE: Amino Acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Linear  
US-08-539-432-17

Query Match 40.0%; Score 6; DB 2; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
|  
Db 1 D 1

RESULT 14  
US-09-327-424-2  
Sequence 2, Application US/09327424  
Patent No. 6027903  
GENERAL INFORMATION:  
APPLICANT: Eyre, David R  
TITLE OF INVENTION: KIT FOR DETECTING ANALYTE INDICATIVE OF TYPE I COLLAGEN  
TITLE OF INVENTION: RESORPTION IN VIVO (as amended)  
FILE REFERENCE: WROS-1-14019  
CURRENT APPLICATION NUMBER: US/09/327,424  
CURRENT FILING DATE: 1993-06-07  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 2  
LENGTH: 2  
TYPE: PRT  
ORGANISM: Human  
FEATURE:  
OTHER INFORMATION: N-telopeptide sequence of type II collagen  
US-09-327-424-2

Query Match 40.0%; Score 6; DB 3; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
|  
Db 2 D 2

RESULT 15  
US-08-480-640A-15  
Sequence 15, Application US/08480640A  
Patent No. 6033904  
GENERAL INFORMATION:  
APPLICANT: Cochran, Mark D.  
APPLICANT: Junker, David E.  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:

ADDRESSEE: John P. White  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,640A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
REGISTRATION NUMBER: 28,678  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-480-640A-15

Query Match 40.0%; Score 6; DB 3; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
|  
Db 1 D 1

Search completed: June 5, 2003, 15:32:43  
Job time : 8.78788 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:31:01 ; Search time 12.1212 Seconds

(Without alignments)  
85.173 Million cell updates/sec

Title: US-09-150-947f-15

Perfect score: 15

Sequence: 1 xxxxxxxxxxxx 10

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCOT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCOTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	40.0	2	9	US-10-187-339-6
2	6	40.0	2	9	US-10-206-699-296
3	6	40.0	2	10	US-09-983-172-8
4	6	40.0	2	10	US-09-983-172-105
5	6	40.0	3	9	US-09-900-936-11
6	6	40.0	3	9	US-09-840-277-31
7	6	40.0	3	9	US-09-747-408-21
8	6	40.0	3	9	US-10-014-162-37
9	6	40.0	3	9	US-10-039-876A-3
10	6	40.0	3	9	US-10-039-876A-4
11	6	40.0	3	9	US-10-015-979-58
12	6	40.0	3	9	US-10-226-956-316
13	6	40.0	3	9	US-10-226-956-317
14	6	40.0	3	9	US-10-226-956-318
15	6	40.0	3	9	US-10-226-956-319
16	6	40.0	3	9	US-10-128-198-1
17	6	40.0	3	9	US-10-237-850-81
18	6	40.0	3	9	US-10-032-818-6
19	6	40.0	3	9	US-10-281-652-11

20	6	40.0	3	9	US-10-283-838-13	Sequence 13, Appl
21	6	40.0	3	9	US-10-163-587A-19	Sequence 19, Appl
22	6	40.0	3	9	US-10-141-531-15	Sequence 15, Appl
23	6	40.0	3	9	US-10-141-531-16	Sequence 16, Appl
24	6	40.0	3	9	US-10-141-531-17	Sequence 17, Appl
25	6	40.0	3	9	US-10-141-531-18	Sequence 18, Appl
26	6	40.0	3	9	US-10-141-531-19	Sequence 19, Appl
27	6	40.0	3	9	US-10-141-531-20	Sequence 20, Appl
28	6	40.0	3	10	US-09-096-749A-113	Sequence 113, App
29	6	40.0	3	10	US-09-866-824A-9	Sequence 9, Appl
30	6	40.0	3	10	US-09-792-200B-19	Sequence 19, Appl
31	6	40.0	3	10	US-09-771-192-11	Sequence 11, Appl
32	6	40.0	3	10	US-09-853-918-11	Sequence 11, Appl
33	6	40.0	3	10	US-09-853-918-6	Sequence 6, Appl
34	6	40.0	3	10	US-09-682-667-14	Sequence 14, Appl
35	6	40.0	3	10	US-09-867-847-30	Sequence 30, Appl
36	6	40.0	3	10	US-09-922-261-7	Sequence 7, Appl
37	6	40.0	3	10	US-09-312-762A-12	Sequence 12, Appl
38	6	40.0	3	10	US-09-982-172-184	Sequence 184, App
39	6	40.0	3	10	US-09-837-697A-11	Sequence 11, Appl
40	6	40.0	3	12	US-10-003-674A-6	Sequence 6, Appl
41	6	40.0	4	1	US-08-464-363-35	Sequence 35, Appl
42	6	40.0	4	1	US-08-610-220A-2	Sequence 2, Appl
43	6	40.0	4	1	US-08-610-220A-7	Sequence 7, Appl
44	6	40.0	4	1	US-08-610-220A-8	Sequence 8, Appl
45	6	40.0	4	8	US-08-765-244-16	Sequence 16, Appl

#### ALIGNMENTS

RESULT 1  
US-10-187-339-6  
; Sequence 6, Application US/10187339  
; Publication No. US20030084478A1  
; GENERAL INFORMATION:  
; APPLICANT: Ream, Walt et al.,  
; TITLE OF INVENTION: Plants Having Enhanced Gall Resistance and  
; TITLE OF INVENTION: Methods and Compositions for Producing the Same  
; FILE REFERENCE: 53629  
; CURRENT APPLICATION NUMBER: US/10/187,339  
; CURRENT FILING DATE: 2002-06-28  
; PRIOR APPLICATION NUMBER: US/09/434,837  
; PRIOR FILING DATE: 1999-11-04  
; PRIOR APPLICATION NUMBER: 60/107,185  
; PRIOR FILING DATE: 1998-11-05  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Agrobacterium tumefaciens  
US-10-187-339-6

Query Match 40.0%; Score 6; DB 9; Length 2;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
DB 2 D 2

RESULT 2  
US-10-206-699-296  
; Sequence 296, Application US/10206699  
; Publication No. US20030100510A1  
; GENERAL INFORMATION:  
; APPLICANT: Sundaramoorthy, M.  
; TITLE OF INVENTION: Crystallized structure of Type IV Collagen NC1 Domain Hexamer  
; FILE REFERENCE: MBHB 01-1017  
; CURRENT APPLICATION NUMBER: US/10/206,699

```

; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/308,523
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/351,289
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 60/366,854
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 60/385,362
; PRIOR FILING DATE: 2002-06-03
; NUMBER OF SEQ ID NOS: 307
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 296
; LENGTH: 2
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-206-699-296

```

```

Query Match          40.0%; Score 6; DB 9; Length 2;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      10 D 10
DB      2 D 2

```

```

RESULT 3
US-09-982-172-8
; Sequence 8, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emli Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THERAGAINS, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 2
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-8

```

```

Query Match          40.0%; Score 6; DB 10; Length 2;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      10 D 10
DB      1 D 1

```

```

RESULT 4
US-09-982-172-105
; Sequence 105, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emli Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THERAGAINS, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 105
; LENGTH: 2

```

```

; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-105

```

```

Query Match          40.0%; Score 6; DB 10; Length 2;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      10 D 10
DB      1 D 1

```

```

RESULT 5
US-09-900-936-11
; Sequence 11, Application US/09900936
; Patent No. US20020165141A1
; GENERAL INFORMATION:
; APPLICANT: dizege, Kathleen
; APPLICANT: Rodgers, Kathleen
; TITLE OF INVENTION: Methods for Promoting Dendritic Cell Proliferation
; FILE REFERENCE: 00-506-A
; CURRENT APPLICATION NUMBER: US/09/900,936
; CURRENT FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:All (1-3)
US-09-900-936-11

```

```

Query Match          40.0%; Score 6; DB 9; Length 3;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      10 D 10
DB      1 D 1

```

```

RESULT 6
US-09-840-277-31
; Sequence 31, Application US/09840277
; Patent No. US20020168363A1
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: KOHNO, TADAHITO
; APPLICANT: LACEY, DAVID LEE
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: INTEGRIN/ADHESION ANTAGONISTS
; FILE REFERENCE: A-688A
; CURRENT APPLICATION NUMBER: US/09/840,277
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 60/198,919
; PRIOR FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 60/201,394
; PRIOR FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Integrin antagonist peptide
US-09-840-277-31

```

Query Match 40.0%; Score 6; DB 9; Length 3;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
DB 3 D 3

## RESULT 7

US-09-747-408-21  
; Sequence 21, Application US/09747408  
; Publication No. US2003003141A1  
; GENERAL INFORMATION:  
; APPLICANT: Green, Allan M.  
; TITLE OF INVENTION: Compounds And Methods For Modulating  
; Cerebral Amyloid Angiopathy  
; FILE REFERENCE: NBI-088  
; CURRENT APPLICATION NUMBER: US/09/747,408  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/171,877  
; PRIOR FILING DATE: 1999-12-23  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 21  
; LENGTH: 3  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-747-408-21

Query Match 40.0%; Score 6; DB 9; Length 3;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
DB 1 D 1

## RESULT 8

US-10-014-162-37  
; Sequence 37, Application US/10014162  
; Publication No. US20030032096A1  
; GENERAL INFORMATION:  
; APPLICANT: Usdin, Ted B.  
; TITLE OF INVENTION: PARATHYROID HORMONE RECEPTOR LIGANDS  
; FILE REFERENCE: NIH/75,001C1  
; CURRENT APPLICATION NUMBER: US/10/014,162  
; PRIOR FILING DATE: 2001-12-11  
; PRIOR APPLICATION NUMBER: PCT/US00/1677  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 60/139335  
; PRIOR FILING DATE: 1999-06-15  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 37  
; LENGTH: 3  
; TYPE: PRT  
; ORGANISM: Bos taurus  
US-10-014-162-37

Query Match 40.0%; Score 6; DB 9; Length 3;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
DB 1 D 1

## RESULT 9

US-10-039-876A-3  
; Sequence 3, Application US/10039876A  
; Publication No. US20030032792A1  
; GENERAL INFORMATION:  
; APPLICANT: Conklin, Darrell C.  
; TITLE OF INVENTION: A HUMAN 2-19 PROTEIN HOMOLOGUE, 2219A  
; FILE REFERENCE: 97-63C1  
; CURRENT APPLICATION NUMBER: US/10/039,876A  
; PRIOR FILING DATE: 2001-10-26  
; PRIOR APPLICATION NUMBER: US 60/061,712  
; PRIOR FILING DATE: 1997-10-06  
; PRIOR APPLICATION NUMBER: US 09/167,513  
; PRIOR FILING DATE: 1998-10-06  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 3  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Motif 1, corresponding to residues 127 to 129 of  
US-10-039-876A-3

Query Match 40.0%; Score 6; DB 9; Length 3;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
DB 2 D 2

## RESULT 10

US-10-039-876A-4  
; Sequence 4, Application US/10039876A  
; Publication No. US20030032792A1  
; GENERAL INFORMATION:  
; APPLICANT: Blumberg, Hal  
; TITLE OF INVENTION: A HUMAN 2-19 PROTEIN HOMOLOGUE, 2219A  
; FILE REFERENCE: 97-63C1  
; CURRENT APPLICATION NUMBER: US/10/039,876A  
; PRIOR FILING DATE: 2001-10-26  
; PRIOR APPLICATION NUMBER: US 60/061,712  
; PRIOR FILING DATE: 1997-10-06  
; PRIOR APPLICATION NUMBER: US 09/167,513  
; PRIOR FILING DATE: 1998-10-06  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 3  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Motif 2, corresponding to residues 156 to 158 of  
US-10-039-876A-4

Query Match 40.0%; Score 6; DB 9; Length 3;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
DB 2 D 2

## RESULT 11

US-10-015-979-58  
; Sequence 58, Application US/10015979  
; Publication No. US20030036050A1

```
;; GENERAL INFORMATION:
;; APPLICANT: Quax, Wilhelmus J.
;; APPLICANT: Verhaert, Raymond M.D.
;; APPLICANT: Beekwilder, Martinus J.
;; APPLICANT: Aehle, Wolfgang
;; TITLE OF INVENTION: Enzyme Selection
;; FILE REFERENCE: 2183-5207US
;; CURRENT APPLICATION NUMBER: US/10/015,979
;; CURRENT FILING DATE: 2001-12-10
;; PRIOR APPLICATION NUMBER: PCT/NL00/00399
;; PRIOR FILING DATE: 2000-06-09
;; PRIOR APPLICATION NUMBER: 60/138,443
;; PRIOR FILING DATE: 1999-06-10
;; NUMBER OF SEQ ID NOS: 99
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 58
;; LENGTH: 3
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: CB4 domain of mutant 13
;; NAME/KEY: SITE
;; LOCATION: (1)-(3)
;; OTHER INFORMATION:
US-10-015-979-58
```

```
Query Match          40.0%; Score 6; DB 9; Length 3;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      10 D 10
DB      3 D 3
```

```
RESULT 12
US-10-226-956-316
;; Sequence 316, Application US/10226956
;; Publication No. US20030060399A1
;; GENERAL INFORMATION:
;; APPLICANT: Brophy, Colleen
;; APPLICANT: Komalavilas, Padmini
;; APPLICANT: Panitch, Alyssa
;; APPLICANT: Joshl, Lokesh
;; APPLICANT: Seal, Brandon L.
;; TITLE OF INVENTION: REAGENTS AND METHODS FOR SMOOTH MUSCLE THERAPIES
;; FILE REFERENCE: ASU-1061-US
;; CURRENT APPLICATION NUMBER: US/10/226,956
;; CURRENT FILING DATE: 2002-08-23
;; PRIOR APPLICATION NUMBER: 60/314,535
;; PRIOR FILING DATE: 2001-08-23
;; NUMBER OF SEQ ID NOS: 320
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 316
;; LENGTH: 3
;; TYPE: PRT
;; ORGANISM: Artificial sequence
;; FEATURE:
;; OTHER INFORMATION: Synthetic peptide
US-10-226-956-316
```

```
Query Match          40.0%; Score 6; DB 9; Length 3;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      10 D 10
DB      1 D 1
```

```
RESULT 13
US-10-226-956-317
```

```
;; Sequence 317, Application US/10226956
;; Publication No. US20030060399A1
;; GENERAL INFORMATION:
;; APPLICANT: Brophy, Colleen
;; APPLICANT: Komalavilas, Padmini
;; APPLICANT: Panitch, Alyssa
;; APPLICANT: Joshl, Lokesh
;; APPLICANT: Seal, Brandon L.
;; TITLE OF INVENTION: REAGENTS AND METHODS FOR SMOOTH MUSCLE THERAPIES
;; FILE REFERENCE: ASU-1061-US
;; CURRENT APPLICATION NUMBER: US/10/226,956
;; CURRENT FILING DATE: 2002-08-23
;; PRIOR APPLICATION NUMBER: 60/314,535
;; PRIOR FILING DATE: 2001-08-23
;; NUMBER OF SEQ ID NOS: 320
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 317
;; LENGTH: 3
;; TYPE: PRT
;; ORGANISM: Artificial sequence
;; FEATURE:
;; OTHER INFORMATION: Synthetic peptide
US-10-226-956-317
```

```
Query Match          40.0%; Score 6; DB 9; Length 3;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      10 D 10
DB      1 D 1
```

```
RESULT 14
US-10-226-956-318
;; Sequence 318, Application US/10226956
;; Publication No. US20030060399A1
;; GENERAL INFORMATION:
;; APPLICANT: Brophy, Colleen
;; APPLICANT: Komalavilas, Padmini
;; APPLICANT: Panitch, Alyssa
;; APPLICANT: Joshl, Lokesh
;; APPLICANT: Seal, Brandon L.
;; TITLE OF INVENTION: REAGENTS AND METHODS FOR SMOOTH MUSCLE THERAPIES
;; FILE REFERENCE: ASU-1061-US
;; CURRENT APPLICATION NUMBER: US/10/226,956
;; CURRENT FILING DATE: 2002-08-23
;; PRIOR APPLICATION NUMBER: 60/314,535
;; PRIOR FILING DATE: 2001-08-23
;; NUMBER OF SEQ ID NOS: 320
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 318
;; LENGTH: 3
;; TYPE: PRT
;; ORGANISM: Artificial sequence
;; FEATURE:
;; OTHER INFORMATION: Synthetic peptide
US-10-226-956-318
```

```
Query Match          40.0%; Score 6; DB 9; Length 3;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      10 D 10
DB      1 D 1
```

```
RESULT 15
US-10-226-956-319
;; Sequence 319, Application US/10226956
;; Publication No. US20030060399A1
;; GENERAL INFORMATION:
```

APPLICANT: Brophy, Colleen  
APPLICANT: Komalavilas, Padmini  
APPLICANT: Panitch, Alyssa  
APPLICANT: Joshi, Lokesh  
APPLICANT: Seal, Brandon L.  
TITLE OF INVENTION: REAGENTS AND METHODS FOR SMOOTH MUSCLE THERAPIES  
FILE REFERENCE: ASU-1061-US  
CURRENT APPLICATION NUMBER: US/10/226,956  
CURRENT FILING DATE: 2002-08-23  
PRIOR APPLICATION NUMBER: 60/314,535  
PRIOR FILING DATE: 2001-08-23  
NUMBER OF SEQ ID NOS: 320  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 319  
LENGTH: 3  
TYPE: PRT  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: Synthetic peptide  
US-10-226-956-319

Query Match 40.0%; Score 6; DB 9; Length 3;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
1  
Db 1 D 1

Search completed: June 5, 2003, 15:49:05  
Job time : 14.1212 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:24:01 ; Search time 9.84848 Seconds  
(without alignments)  
97.613 Million cell updates/sec

Title: US-09-150-947F-15

Perfect score: 15

Sequence: 1 XXXXXXXXXXD 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: PIR\_73:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	40.0	4	2	I40697
2	6	40.0	4	2	A41890
3	6	40.0	4	2	D41654
4	6	40.0	4	2	A32480
5	6	40.0	4	2	I51049
6	6	40.0	4	2	PT0271
7	6	40.0	4	2	B53284
8	6	40.0	4	2	PT0696
9	6	40.0	4	2	PT0711
10	6	40.0	4	2	A26209
11	6	40.0	5	2	A32516
12	6	40.0	5	2	C23751
13	6	40.0	5	2	A26830
14	6	40.0	5	2	A32014
15	6	40.0	5	2	B31836
16	6	40.0	5	2	B60274
17	6	40.0	5	2	D60274
18	6	40.0	5	2	S70615
19	6	40.0	5	2	PQ0689
20	6	40.0	5	2	PT0267
21	6	40.0	5	2	PT0281
22	6	40.0	5	2	PT0308
23	6	40.0	5	2	PT0536
24	6	40.0	5	2	PT0513
25	6	40.0	5	2	PT0600
26	6	40.0	5	2	PT0729
27	6	40.0	5	2	PT0624
28	6	40.0	5	2	PT0601
29	6	40.0	5	2	PT0672

30	6	40.0	5	2	PT0660	T-cell receptor be
31	6	40.0	5	2	PT0651	T-cell receptor be
32	6	40.0	5	2	PT0656	T-cell receptor be
33	6	40.0	5	2	PT0535	T-cell receptor be
34	6	40.0	5	2	PT0639	T-cell receptor be
35	6	40.0	5	2	PT0538	T-cell receptor be
36	6	40.0	5	2	PT0561	T-cell receptor be
37	6	40.0	5	2	PT0540	T-cell receptor be
38	6	40.0	5	2	PT0703	T-cell receptor be
39	6	40.0	5	2	PT0690	T-cell receptor be
40	6	40.0	5	2	PT0573	T-cell receptor be
41	6	40.0	5	2	PT0580	T-cell receptor be
42	6	40.0	5	2	PT0679	T-cell receptor be
43	6	40.0	5	2	S68326	blood cell protein
44	6	40.0	6	2	A61419	sarcosine dehydrog
45	6	40.0	6	2	B44510	hypothetical prote

#### ALIGNMENTS

RESULT 1  
I40697  
biotin A - Citrobacter freundii (fragment)  
C/Species: Citrobacter freundii  
C/Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 12-Aug-1996  
C/Accession: I40697  
R/Shiuan, D.; Campbell, A.  
Gene 67, 203-211, 1988  
A/Title: Transcriptional regulation and gene arrangement of Escherichia coli, Citrobacter  
A/Reference number: I40697; MVID:89006280; PMID:2971595  
A/Accession: I40697  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 14 <RES>  
A/Cross-references: GB:M21922; NID:9144434

Query Match 40.0%; Score 6; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
Db 3 D 3

#### RESULT 2

A41890  
protein D - Escherichia coli (fragment)  
C/Species: Escherichia coli  
C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 30-Sep-1993  
C/Accession: A41890  
R/Stettan, A.; Gebhardt, K.; Kristiansen, E.; Birreland, N.K.; Lindqvist, B.H.  
J. Bacteriol. 174, 4094-4100, 1992  
A/Title: Escherichia coli K-12 and B contain functional bacteriophage P2 ogf genes.  
A/Reference number: A41890; MVID:92283767; PMID:1597424  
A/Accession: A41890  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-4 <SLE>  
A/Cross-references: GB:M81463

Query Match 40.0%; Score 6; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
Db 2 D 2

#### RESULT 3

D41654

hypothetical protein (sodc 5' region) - Haemophilus parainfluenzae (fragment)  
C:Species: Haemophilus parainfluenzae  
C:Date: 12-Jun-1992 #sequence\_revision 12-Jun-1992 #text\_change 24-Feb-1995  
C:Accession: D41654

R:Kroll, J.S.; Langford, P.R.; Loynds, B.M.  
J. Bacteriol. 173, 7449-7457, 1991

A:Title: Copper-zinc superoxide dismutase of Haemophilus influenzae and Haemophilus parainfluenzae  
A:Reference number: A1654; MUID:92041655; PMID:1938942

A:Accession: D41654

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-4 <KRO>

Query Match

Best Local Similarity 40.0%; Score 6; DB 2; Length 4;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10

DB 1 D 1

RESULT 4

A32480

A:Title: Achatin-II - giant African snail

N:Contains: Achatin-II

C:Species: Achatina fulica (giant African snail)

C:Date: 12-Oct-1989 #sequence\_revision 12-Oct-1989 #text\_change 17-Mar-1999

C:Accession: A32480

R:Kamatani, Y.; Minakata, H.; Kenny, P.T.M.; Iwashita, T.; Watanabe, K.; Funase, K.; Sun

Biochem. Biophys. Res. Commun. 160, 1015-1020, 1989

A:Title: Achatin-II, an endogenous neuroexcitatory tetrapeptide from Achatina fulica feru

A:Reference number: A32480; MUID:89273551; PMID:2597281

A:Accession: A32480

A:Molecule type: Protein

A:Residues: 1-4 <KAM>

A:Note: stereochemistry of the active form confirmed by chemical synthesis

R:Ishida, T.; In, Y.; Inoue, M.; Yasuda-Kamatani, Y.; Minakata, H.; Iwashita, T.; Nomoto

FEBS Lett. 307, 253-256, 1992

A:Title: Effect of the D-Phe(2) residue on molecular conformation of an endogenous neuro

(H-Gly-Phe-Ala-Asp-OH).  
A:Reference number: A44691; MUID:92354723; PMID:1644179

A:Contents: annotation: X-ray crystallography, 0.85 angstroms

A:Note: Achatin-II has L-phenylalanine

C:Keywords: D-amino acid

F:2/Modified site: D-phenylalanine (Phe) #status experimental

Query Match

Best Local Similarity 40.0%; Score 6; DB 2; Length 4;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10

DB 4 D 4

RESULT 5

I51049

A:Title: metallochionein-A - rainbow trout (fragment)

C:Species: Oncorhynchus mykiss (rainbow trout)

C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000

C:Accession: I51049

R:Olsson, P.E.; Kiling, P.; Erkel, L.J.; Kille, P.

Eur. J. Biochem. 230, 344-349, 1995

A:Title: Structural and functional analysis of the rainbow trout (Oncorhynchus mykiss) me

A:Reference number: I51049; MUID:95324545; PMID:7601121

A:Accession: I51049

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-4 <OIS>

A:Cross-references: EMBL:X80181; NID:91019799; PIDN:CA56466.1; PID:94379328

Query Match

40.0%; Score 6; DB 2; Length 4;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10

DB 2 D 2

RESULT 6

PT0271

Ig heavy chain CRD3 region (clone 3-103A) - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996

C:Accession: PT0271

R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.

J. Exp. Med. 173, 395-407, 1991

A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity an

A:Reference number: PT0222; MUID:91108337; PMID:1899102

A:Accession: PT0271

A:Molecule type: DNA

A:Residues: 1-4 <YAM>

A:Experimental source: B lymphocyte

C:Keywords: heterotetramer; immunoglobulin

Query Match

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10

DB 1 D 1

RESULT 7

B53284

T-cell receptor beta 2 chain D region, Dbeta2 - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 02-May-1994 #sequence\_revision 18-Nov-1994 #text\_change 05-Nov-1999

C:Accession: B53284

R:Harindranath, N.; Alexander, C.B.; Mage, R.G.

Mol. Immunol. 28, 881-888, 1991

A:Title: Evolutionarily conserved organization and sequences of germline diversity an

A:Reference number: A53284; MUID:91342695; PMID:1678859

A:Accession: B53284

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-4 <HAR>

A:Cross-references: GB:S60737; NID:9233916; PIDN:AB19518.1; PID:9233918

A:Note: sequence extracted from NCBI Backbone (NCBIN:60737; NCBIP:60738)

C:Keywords: T-cell receptor

Query Match

Best Local Similarity 40.0%; Score 6; DB 2; Length 4;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10

DB 1 D 1

RESULT 8

PT0696

T-cell receptor beta chain V-D-J region (100-2N) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997

C:Accession: PT0696; PT0612; PT0545; PT0692; PT0552; PT0696

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions

A:Reference number: PT0509; MUID:91277601; PMID:1711558

A:Accession: PT0696

A:Status: translation not shown

A:Molecule type: mRNA



A:Residues: 1-4 <FE2>  
 A:Experimental source: newborn thymus, strain BALB/c, 100-2M  
 A:Accession: PT0612  
 A:Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-4 <FE2>  
 A:Experimental source: newborn thymus, strain BALB/c, 111-1J  
 A:Accession: PT0545  
 A:Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-4 <FE3>  
 A:Experimental source: day 18 fetal thymus, strain BALB/c, 126-1AD  
 A:Accession: PT0692  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-4 <FE4>  
 A:Experimental source: day 18 fetal thymus, strain BALB/c, 140-1M  
 A:Accession: PT0552  
 A:Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-4 <FE5>  
 A:Experimental source: day 18 fetal thymus, strain BALB/c, 126-1CI  
 A:Accession: PT0696  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-4 <FE6>  
 A:Experimental source: newborn thymus, strain BALB/c, 135-1AA  
 C:Keywords: T-cell receptor

Query Match 40.0%; Score 6; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
 |  
 Db 4 D 4

## RESULT 9

PT0711  
 T-cell receptor beta chain V-D-J region (120-2J) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C>Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
 C:Accession: PT0607; PT0674; PT0570; PT0711; PT0710  
 R:Feeney, A.J.  
 J:Exp. Med. 174, 115-124, 1991  
 A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
 A:Reference number: PT0509; MUID:91277601; PMID:1711558  
 A:Accession: PT0607  
 A:Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-4 <FE1>  
 A:Experimental source: newborn thymus, strain BALB/c, 120-2J  
 A:Accession: PT0674  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-4 <FE2>  
 A:Experimental source: day 18 fetal thymus, strain BALB/c, 140-1G  
 A:Accession: PT0678  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-4 <FE3>  
 A:Experimental source: day 18 fetal thymus, strain BALB/c, 154-1L  
 A:Accession: PT0570  
 A:Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-4 <FE4>  
 A:Experimental source: day 19 fetal thymus, strain BALB/c, 141-1I  
 A:Accession: PT0711  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-4 <FE5>  
 A:Experimental source: newborn thymus, strain BALB/c (clones 161-2AE and 161-2AF)

C:Keywords: T-cell receptor

Query Match 40.0%; Score 6; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
 |  
 Db 4 D 4

## RESULT 10

A26209  
 protein-glutamine gamma-glutamyltransferase (EC 2.3.2.13) - guinea pig (fragment)  
 C:Species: Cavia porcellus (guinea pig)  
 C>Date: 10-Sep-1987 #sequence\_revision 10-Sep-1987 #text\_change 03-Mar-1995  
 C:Accession: A26209  
 R:Connellan, J.M.; Chung, S.I.; Whetzel, N.K.; Bradley, L.M.; Folk, J.E.  
 J: Biol. Chem. 246, 1093-1098, 1971  
 A:Title: Structural properties of guinea pig liver transglutaminase.  
 A:Reference number: A26209; MUID:7111415; PMID:5543674  
 A:Accession: A26209  
 A:Molecule type: protein  
 A:Residues: 1-4 <CON>  
 A:Experimental source: liver  
 C:Keywords: aminoacyltransferase

Query Match 40.0%; Score 6; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
 |  
 Db 3 D 3

## RESULT 11

A32516  
 cholecystokinin-5 - dog  
 N:Alternate names: CCK-5  
 C:Species: Canis lupus familiaris (dog)  
 C>Date: 18-Oct-1989 #sequence\_revision 18-Oct-1989 #text\_change 18-Aug-2000  
 C:Accession: A32516  
 R:Shively, J.; Reeve Jr., J.R.; Eysselein, V.E.; Ben-Avram, C.; Vigna, S.R.; Walsh, J.  
 Am. J. Physiol. 252, G272-G275, 1987  
 A:Title: CCK-5: sequence analysis of a small cholecystokinin from canine brain and in  
 A:Reference number: A32516; MUID:87153871; PMID:3626554  
 A:Accession: A32516  
 A:Molecule type: protein  
 A:Residues: 1-5 <SH1>  
 C:Comment: This peptide corresponds to the five carboxyl-terminal residues of cholecy  
 C:Superfamily: gastrin  
 C:Keywords: amidated carboxyl end; neuropeptide  
 F:5/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 40.0%; Score 6; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
 |  
 Db 4 D 4

## RESULT 12

C23751  
 spinal cord peptide SCP-6 - pig  
 C:Species: Sus scrofa domestica (domestic pig)  
 C>Date: 28-Sep-1987 #sequence\_revision 28-Sep-1987 #text\_change 16-Aug-2000  
 C:Accession: C23751  
 R:Hsi, K.L.; Chen, R.L.; Chen, Z.G.; Zhang, H.L.; Lu, Y.A.; Guo, S.Y.; Wu, S.X.; Tsou  
 Arch. Biochem. Biophys. 240, 178-183, 1985  
 A:Reference number: A23751; MUID:85250425; PMID:4015098

A:Accession: C23751  
A:Status: preliminary  
A:Molecule type: Protein  
A:Residues: 1-5 <HS1>  
C:Superfamily: unassigned animal peptides

Query Match 40.0%; Score 6; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 10 D 10  
|  
Db 1 D 1

## RESULT 13

A26830  
mitosis inhibiting peptide - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 19-Nov-1988 #sequence\_revision 18-Aug-2000 #text\_change 18-Aug-2000  
C:Accession: A26830  
R:Reichelt, K.; Eljjo, K.; Edmonson, P.D.  
Biochem. Biophys. Res. Commun. 146, 1493-1501, 1987  
A:Title: Isolation and structure of an epidermal mitosis inhibiting pentapeptide.  
A:Reference number: A26830; MUID:87298602; PMID:3619940  
A:Accession: A26830  
A:Molecule type: Protein  
A:Residues: 1-5 <REI>  
C:Superfamily: unassigned animal peptides  
C:Keywords: blocked amino end; pyroglutamic acid  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 40.0%; Score 6; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
|  
Db 3 D 3

## RESULT 14

A32014  
tram protein - Escherichia coli plasmid R100 (fragment)  
C:Species: Escherichia coli  
C:Date: 22-Jun-1989 #sequence\_revision 22-Jun-1989 #text\_change 16-Feb-1997  
C:Accession: A32014  
R:Inamoto, S.; Yoshio, Y.; Ohtsubo, E.  
J. Bacteriol. 170, 2749-2757, 1988  
A:Title: Identification and characterization of the products from the traJ and traY gene  
A:Reference number: A32014; MUID:88227859; PMID:2836369  
A:Accession: A32014  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-5 <TNA>  
C:Genetics:  
A:Genome: plasmid  
C:Keywords: DNA binding

Query Match 40.0%; Score 6; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
|  
Db 3 D 3

## RESULT 15

B31836  
20K proteoln - Rickettsia rickettsii (fragment)  
C:Species: Rickettsia rickettsii  
C:Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 28-May-1999

C:Accession: B31836  
R:Anderson, B.E.; Baumstark, B.R.; Bellini, W.J.  
J. Bacteriol. 170, 4493-4500, 1988  
A:Title: Expression of the gene encoding the 17-kilodalton antigen from Rickettsia r  
A:Reference number: A31836; MUID:89008059; PMID:3139629  
A:Accession: B31836  
A:Molecule type: DNA  
A:Residues: 1-5 <AND>  
A:Cross-references: GB:J03371; NID:9152455; PID:AA015030.1; PID:94262874

Query Match 40.0%; Score 6; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
|  
Db 2 D 2

Search completed: June 5, 2003, 15:34:00  
Job time : 10.8485 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:34:07 ; Search time 6.5155 Seconds

(without alignments)  
63.661 Million cell updates/sec

Title: US-09-150-947F-15  
Perfect score: 15  
Sequence: 1 XXXXXXXXXXD 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues  
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	6	40.0	3	1	LUXE_VIBFI
2	6	40.0	4	1	ACH1_ACHFU
3	6	40.0	4	1	OCPI_OCTMI
4	6	40.0	4	1	OCPI_OCTMI
5	6	40.0	5	1	BIOA_CITFR
6	6	40.0	5	1	RELI_LITRU
7	6	40.0	5	1	TRM3_ECOLI
8	6	40.0	5	1	UXM4_CHUTR
9	6	40.0	6	1	ASP2_LACSN
10	6	40.0	6	1	TRPI_PSEPU
11	6	40.0	7	1	ALL4_CARMA
12	6	40.0	7	1	ALL7_CYDPO
13	6	40.0	7	1	CHOK_AICSP
14	6	40.0	7	1	PAR1_HELTI
15	6	40.0	7	1	PAR2_PROCL
16	6	40.0	7	1	FAR2_CALVO
17	6	40.0	7	1	UF03_MOUSE
18	6	40.0	7	1	UF04_MOUSE
19	6	40.0	8	1	ACL_THRAL
20	6	40.0	8	1	ACT_CARMA
21	6	40.0	8	1	AKH_MELML
22	6	40.0	8	1	AL12_CARMA
23	6	40.0	8	1	AL18_CARMA
24	6	40.0	8	1	ALL4_CALVO
25	6	40.0	8	1	ALL5_CALVO
26	6	40.0	8	1	ALL5_CYDPO
27	6	40.0	8	1	ANG2_BOTJU
28	6	40.0	8	1	CCRN_MACEU
29	6	40.0	8	1	COXG_RAT
30	6	40.0	8	1	FAR1_PANRE
31	6	40.0	8	1	FAR2_HOMAM
32	6	40.0	8	1	FAR3_CALVO
33	6	40.0	8	1	GLUR_HUMAN

34	6	40.0	8	1	LCK1_LEUMA	P21140 leucophaea
35	6	40.0	8	1	LCK2_LEUMA	P21141 leucophaea
36	6	40.0	8	1	LCK3_LEUMA	P21142 leucophaea
37	6	40.0	8	1	LCK4_LEUMA	P21143 leucophaea
38	6	40.0	8	1	LCK7_LEUMA	P19989 leucophaea
39	6	40.0	8	1	LCK8_LEUMA	P19990 leucophaea
40	6	40.0	8	1	LMT2_LOCLI	P22396 locusta mly
41	6	40.0	8	1	ORMY_ORCLI	P82455 orconectes
42	6	40.0	8	1	PLP_BRANA	P81707 brassica na
43	6	40.0	8	1	UC26_MAIZE	P80632 zea mays (m
44	6	40.0	8	1	UPA1_HUMAN	P30087 homo sapien
45	6	40.0	9	1	COM_CONVE	P83047 conus ventr

## ALIGNMENTS

RESULT 1	ID	Score	Length	ID	Description
LUXE_VIBFI	1	6	40.0	8	1
LUXE_VIBFI	1	6	40.0	8	1
AC P24272	1	6	40.0	8	1
DT 01-MAR-1992 (Rel. 21, Last Created)	1	6	40.0	8	1
DT 01-MAR-1992 (Rel. 21, Last Created)	1	6	40.0	8	1
DT 01-OCT-1996 (Rel. 34, Last annotation update)	1	6	40.0	8	1
DE Long-chain-fatty-acid--luciferin-component ligase (EC 6.2.1.19) (Acyl)-protein synthetase (Fragment).	1	6	40.0	8	1
GN LUXE.	1	6	40.0	8	1
OS Vibrio fischeri.	1	6	40.0	8	1
OC Bacteria: Proteobacteria: gamma subdivision: Vibrionaceae: Vibrio.	1	6	40.0	8	1
OX NCBI_TaxID=668;	1	6	40.0	8	1
RN [1]	1	6	40.0	8	1
RP SEQUENCE FROM N.A.	1	6	40.0	8	1
RX MEDLINE-91072226; PubMed-2254256;	1	6	40.0	8	1
RA Swartzman E., Kapoor S., Graham A.F., Melyghen E.A.:	1	6	40.0	8	1
RT A new Vibrio fischeri lux gene precedes a bidirectional termination site for the lux operon.	1	6	40.0	8	1
RL J. Bacteriol. 172:6797-6802(1990).	1	6	40.0	8	1
CC -1- FUNCTION: ACYL-PROTEIN SYNTHETASE ACTIVATES TETRADECANOIC ACID. IT IS A COMPONENT OF THE FATTY ACID REDUCTASE COMPLEX RESPONSIBLE FOR CONVERTING TETRADECANOIC ACID TO THE ALDEHYDE WHICH SERVES AS SUBSTRATE IN THE LUCIFERASE-CATALYZED REACTION.	1	6	40.0	8	1
CC -1- CATALYTIC ACTIVITY: ATP + an acid + protein - AMP + diphosphate + an acyl-protein thioester.	1	6	40.0	8	1
CC -1- PATHWAY: Bioluminescent fatty acid reduction system; second step.	1	6	40.0	8	1
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).	1	6	40.0	8	1
CC EMBL: M62812; -; NOT_ANNOTATED_CDS.	1	6	40.0	8	1
DR Luminescence; Ligase.	1	6	40.0	8	1
KW NON_TER	1	6	40.0	8	1
FT SEQUENCE 3 AA: 374 MM: 6AA3303000000000 CRC64;	1	6	40.0	8	1
SO SEQUENCE 3 AA: 374 MM: 6AA3303000000000 CRC64;	1	6	40.0	8	1
Query Match	1	6	40.0	8	1
Best Local Similarity	1	6	40.0	8	1
Matches	1	6	40.0	8	1
1; Conservative	1	6	40.0	8	1
0; Mismatches	1	6	40.0	8	1
0; Indels	1	6	40.0	8	1
0; Gaps	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1
0; Score 6;	1	6	40.0	8	1
0; DB 1;	1	6	40.0	8	1
0; Length 3;	1	6	40.0	8	1
0; Pred. No. 0;	1	6	40.0	8	1</

DT 15-JUN-1998 (Rel. 36, Last annotation update)  
 DE Achatin-I.  
 OS Achatina fulica (Giant African snail).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Stylommatophora;  
 CC Achatinacea; Achatinidae; Achatina.  
 OX NCBI\_TaxID=6530;  
 RN [1]  
 RP SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.  
 RC STRAIN-Ferussac; TISSUE-Ganglion;  
 RX MEDLINE=89273551; PubMed=2597281;  
 RA Kametani Y., Minakata H., Kenny P.T.M., Iwashita T., Watanabe K.,  
 RA Funase K., Sun X.P., Yongsiri A., Kim K.H., Novales-Ll P.,  
 RA Novales E.T., Kanapi C.G., Takeuchi H., Nomoto K.,  
 RT "Achatin-I, an endogenous neuroexcitatory tetrapeptide from Achatina  
 fulica ferussac containing a D-amino acid residue."  
 RL Biochem. Biophys. Res. Commun. 160:1015-1020(1989).  
 RN [2]  
 RP CHARACTERIZATION.  
 RC STRAIN-Ferussac; TISSUE-Heart atrium;  
 RX MEDLINE=91264856; PubMed=1675568;  
 RA Fujimoto K., Kubota I., Yasuda-Kamatani Y., Minakata H., Nomoto K.,  
 RA Yoshida M., Harada A., Muneoka Y., Kobayashi M.,  
 RT "Purification of achatin-I from the atria of the African giant snail,  
 Achatina fulica, and its possible function."  
 RL Biochem. Biophys. Res. Commun. 177:847-853(1991).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY.  
 RX MEDLINE=93014529; PubMed=1399265;  
 RA Ishida T., In Y., Doi M., Inoue M., Yasuda-Kamatani Y., Minakata H.,  
 RA Iwashita T., Nomoto K.,  
 RT "Crystal structure and molecular conformation of achatin-I  
 (H-Gly-D-Phe-Ala-Asp-OH), an endogenous neuropeptide containing a  
 D-amino acid residue."  
 RL Int. J. Pept. Protein Res. 39:258-264(1992).  
 CC -1- FUNCTION: NEUROEXCITATORY PEPTIDE; INCREASES THE IMPULSE FREQUENCY  
 AND PRODUCES A SPIKE BROADENING OF THE IDENTIFIED HEART EXCITATORY  
 NEURON (PON); ALSO ENHANCES THE AMPLITUDE AND FREQUENCY OF THE  
 HEART BEAT. HAS ALSO AN EFFECT ON SEVERAL OTHER MUSCLES.  
 DR PIR: A32460; A32480.  
 KW Hormone; D-amino acid.  
 FT MOD\_RES 2  
 SQ SEQUENCE 4 AA; 408 MW; 6AADD9C81000000 CRC64;  
 D-PHENYLANILINE.  
 Query Match 40.0%; Score 6; DB 1; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 10 D 10  
 DB 4 D 4  
 RESULT 3  
 OCP1\_OCTMI STANDARD; PRT: 4 AA.  
 AC P58648;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Cardioactive peptides Ocp-1/Ocp-2.  
 OS Octopus minor (Octopus).  
 CC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;  
 CC Incurtata; Octopodidae; Octopus.  
 OX NCBI\_TaxID=89766;  
 RN [1]  
 RP SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.  
 RC TISSUE-Brain;  
 RX MEDLINE=20336815; PubMed=10876044;  
 RA Iwakoshi E., Hisada M., Minakata H.,  
 RT "Cardioactive peptides isolated from the brain of a Japanese octopus,  
 Octopus minor."  
 RL Peptides 21:623-630(2000).  
 CC -1- FUNCTION: Cardioactive; has both positive chronotropic and  
 inotropic effects on the heart. Ocp-4 is a 1000 time less  
 active than Ocp-3.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- PTM: Ocp-4 has D-Ser instead of L-Ser.  
 CC -1- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.  
 KW Hormone; D-amino acid.  
 FT MOD\_RES 2  
 SQ SEQUENCE 4 AA; 463 MW; 6AB365B81000000 CRC64;  
 D-SERINE (IN OCP-4).  
 Query Match 40.0%; Score 6; DB 1; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 10 D 10  
 DB 4 D 4

CC Inotropic effects on the heart. Ocp-2 is a 1000 time less  
 active than Ocp-1.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- PTM: Ocp-2 has L-Phe instead of D-Phe.  
 CC -1- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.  
 KW Hormone; D-amino acid.  
 FT MOD\_RES 2  
 SQ SEQUENCE 4 AA; 394 MW; 6AA879C81000000 CRC64;  
 D-PHENYLANILINE.  
 Query Match 40.0%; Score 6; DB 1; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 10 D 10  
 DB 4 D 4  
 RESULT 4  
 OCP3\_OCTMI STANDARD; PRT: 4 AA.  
 AC P58649;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Cardioactive peptides Ocp-3/Ocp-4.  
 OS Octopus minor (Octopus).  
 CC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;  
 CC Incurtata; Octopodidae; Octopus.  
 OX NCBI\_TaxID=89766;  
 RN [1]  
 RP SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.  
 RC TISSUE-Brain;  
 RX MEDLINE=20336815; PubMed=10876044;  
 RA Iwakoshi E., Hisada M., Minakata H.,  
 RT "Cardioactive peptides isolated from the brain of a Japanese octopus,  
 Octopus minor."  
 RL Peptides 21:623-630(2000).  
 CC -1- FUNCTION: Cardioactive; has both positive chronotropic and  
 inotropic effects on the heart. Ocp-4 is a 1000 time less  
 active than Ocp-3.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- PTM: Ocp-4 has D-Ser instead of L-Ser.  
 CC -1- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.  
 KW Hormone; D-amino acid.  
 FT MOD\_RES 2  
 SQ SEQUENCE 4 AA; 463 MW; 6AB365B81000000 CRC64;  
 D-SERINE (IN OCP-4).  
 Query Match 40.0%; Score 6; DB 1; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 10 D 10  
 DB 4 D 4  
 RESULT 5  
 BIOA\_CITFR STANDARD; PRT: 5 AA.  
 AC P13071;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Adenosylmethionine-8-amino-7-oxononanoate aminotransferase  
 (EC 2.6.1.62) (7,8-diamino-pelargonic acid aminotransferase) (DAPA  
 aminotransferase) (Fragment).  
 DE BIOA.  
 GN Citrobacter freundii.  
 OS Citrobacter; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 CC Bacteria.  
 OX NCBI\_TaxID=546;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE-89006280; PubMed-2971595;  
 RA Shihuan D., Campbell A.;  
 RT "Transcriptional regulation and gene arrangement of *Escherichia coli*,  
 RT *Citrobacter freundii* and *Salmonella typhimurium* biotin operons.",  
 RL Gene 67:203-211(1988).  
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 8-amino-7-  
 CC oxomalonate -> S-adenosyl-4-methylthio-2-oxobutanoate + 7,8-  
 CC diaminononanoate.  
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.  
 CC -1- PATHWAY: Biotin biosynthesis.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT  
 CC AMINOTRANSFERASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M21922; -; NOT\_ANNOTATED\_CDS.  
 DR Interpro: IPR000934; AminoTran\_3.  
 DR PROSITE: PS00600; AA\_TRANSF\_CLASS\_3; PARTIAL.  
 KM Biotin biosynthesis; Transferase; Amino transferase;  
 KM Pyridoxal phosphate. 5  
 FT NON\_TER  
 SQ SEQUENCE 5 AA: 582 MW: 6AABAB1B1A6F00000 CRC64;  
 QY  
 DB 10 D 10  
 DB 4 D 4  
 Query Match 40.0%; Score 6; DB 1; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 10 D 10  
 DB 4 D 4  
 RESULT 6  
 REEL\_LITRU STANDARD; PRT; 5 AA.  
 AC P82070;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Rubellidin 1.1.  
 OS Litoria rubella (Desert tree frog).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;  
 OC Pelodytidae; Litoria.  
 OC NCBI\_TaxID=104895;  
 RN [1]  
 RP SEQUENCE, AND MASS SPECTROMETRY.  
 RC TISSUE-Skin secretion;  
 RA Steinboerner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,  
 RA Tyler M.J., Wallace J.C.;  
 RT "The structure of new peptides from the Australian red tree frog  
 RT *Litoria rubella*, the skin peptide profile as a probe for the study  
 RT of evolutionary trends of amphibians.",  
 RL Aust. J. Chem. 49:955-963(1996).  
 CC -1- FUNCTION: SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR ANTIBIOTIC  
 CC ACTIVITY.  
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.  
 CC -1- MASS SPECTROMETRY: MW=598; METHOD-FAB.  
 KW Amphibian skin.  
 SQ SEQUENCE 5 AA: 598 MW: 6DD9C9CAB2A00000 CRC64;  
 QY  
 DB 10 D 10  
 DB 4 D 4  
 Query Match 40.0%; Score 6; DB 1; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
 DB 2 D 2  
 RESULT 7  
 TRM3\_ECOLI STANDARD; PRT; 5 AA.  
 AC P13973;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Trm protein (Fragment).  
 OS *Escherichia coli*.  
 GN Plasmid Incipit R100.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC *Escherichia*.  
 OC NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-88227859; PubMed-2836369;  
 RA Inamoto S., Yoshitaka Y., Ohtsuda E.;  
 RT "Identification and characterization of the products from the trm  
 RT and trm genes of plasmid R100.",  
 RL J. Bacteriol. 170:2749-2757(1988).  
 CC -1- FUNCTION: TRANSFER GENE PROTEIN. IS INVOLVED IN THE CONJUGATION  
 CC PROCESS OF BACTERIAL CELLS FOR THE EXCHANGE OF PLASMID DNA.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: TO TRM PROTEIN OF OTHER PLASMIDS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M20941; -; NOT\_ANNOTATED\_CDS.  
 DR PIR: A32014; A32014.  
 KM Conjugation; Plasmid; DNA-binding.  
 FT NON\_TER  
 SQ SEQUENCE 5 AA: 634 MW: 6B1B1AA443500000 CRC64;  
 QY 10 D 10  
 DB 3 D 3  
 Query Match 40.0%; Score 6; DB 1; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 10 D 10  
 DB 3 D 3  
 RESULT 8  
 UX4A\_CHLTR STANDARD; PRT; 5 AA.  
 ID UX4A\_CHLTR  
 AC P38005;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Unknown protein from 2D-page from elementary body (Fragment).  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OC NCBI\_TaxID=813;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN-12/434/Bu;  
 RA Binl E., Santucci A., Magl B., Marzocchi B., Sanchez-Campillo M.,  
 RA Comanducci M., Christensen G., Birkelund S., Vitreou E., Ratli G.,  
 RA Pallini V.;  
 RL Submitted (SEP-1994) to the SWISS-PROT data bank.  
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN

CC PROTEIN IS: 4.5, ITS MW IS: 28 KDa.  
 DR Stena-2DPAGE: P38005; -.  
 FT NON\_TER 5  
 SQ SEQUENCE 5 AA; 474 MW; 75BAA865AA800000 CRC64;  
 Query Match 40.0%; Score 6; DB 1; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 10 D 10  
 DB 4 D 4

RESULT 9  
 ASP2\_LACSN STANDARD; PRT: 6 AA.  
 ID ASP2\_LACSN  
 AC P82655;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Acid shock protein 2 (Fragment).  
 OS Lactobacillus sanfranciscensis (Lactobacillus sanfrancisco).  
 CC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;  
 CC Lactobacillus.  
 CC NCBI\_TaxID=1625;  
 RN (1)  
 RP SEQUENCE.  
 RC STRAIN=CBI;  
 RX MEDLINE-21322712; PubMed-11429463;  
 RA De Angelis M., Bini L., Pallini V., Coccconcelli P.S., Gobetti M.;  
 RT "The acid-stress response in Lactobacillus sanfranciscensis CBI.";  
 RL Microbiology 147:1863-1873(2001).  
 CC -1- INDUCTION: OVEREXPRESSED IN ACID ENVIRONMENTS.  
 FT NON\_TER 6  
 SQ SEQUENCE 6 AA; 778 MW; 6AA45B5B132AB000 CRC64;  
 Query Match 40.0%; Score 6; DB 1; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 10 D 10  
 DB 6 D 6

RESULT 10  
 TRP1\_PSEPU STANDARD; PRT: 6 AA.  
 ID TRP1\_PSEPU  
 AC P36414;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-FEB-1995 (Rel. 31, Last annotation update)  
 DE TRPBA operon transcriptional activator (Fragment).  
 GN TRP1.  
 OS Pseudomonas putida.  
 CC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 CC Pseudomonas.  
 CC NCBI\_TaxID=303;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PEL C15;  
 RX MEDLINE-89335826; PubMed-2503057;  
 RA Ederly L., Crawford I.P.;  
 RT "DNA sequence of the tryptophan synthase genes of Pseudomonas putida";  
 RL Biochimie 71:521-531(1989).  
 CC -1- FUNCTION: ACTIVATES THE EXPRESSION OF THE TRPA GENES ENCODING THE TWO TRYPTOPHAN SYNTHASE SUBUNITS. IN THE ABSENCE OF THE INDUCER (INDOLEGLYCEROL PHOSPHATE), TRP1 BINDS UPSTREAM OF THE TRPA OPERON, OVERLAPPING ITS OWN PROMOTER REGION.  
 CC -1- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL REGULATORS.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC EMBL: X13299; CAA31660.1; -.  
 DR InterPro: IPR000847; HTH\_LYSR.  
 DR PROSITE: PS00044; HTH\_LYSR\_FAMILY, PARTIAL.  
 KW Tryptophan biosynthesis; Transcription regulation; Activator;  
 KW DNA-binding.  
 FT NON\_TER 6  
 SQ SEQUENCE 6 AA; 683 MW; 77672AA1EDD6F000 CRC64;  
 Query Match 40.0%; Score 6; DB 1; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 10 D 10  
 DB 4 D 4

RESULT 11  
 AL14\_CARMA STANDARD; PRT: 7 AA.  
 ID AL14\_CARMA  
 AC P81807;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Carinustatin 4.  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;  
 CC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;  
 CC Brachyura; Eubrachyura; Portunoidae; Portunidae; Carcinus.  
 CC NCBI\_TaxID=6759;  
 RN (1)  
 RP SEQUENCE.  
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;  
 RX MEDLINE-98121193; PubMed-9461295;  
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
 RA Thorpe A.;  
 RT "Isolation and identification of multiple neuropeptides of the allatostatin superfamily in the shore crab Carcinus maenas";  
 RL Eur. J. Biochem. 250:727-734(1997).  
 CC -1- FUNCTION: MAY ACT AS A NEUROTROPHIC OR NEUROMODULATOR.  
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 CC Neuropeptide; Multigene family.  
 SQ SEQUENCE 7 AA; 782 MW; 672879CDB476AC0 CRC64;  
 Query Match 40.0%; Score 6; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 10 D 10  
 DB 1 D 1

RESULT 12  
 ALL7\_CYPDPO STANDARD; PRT: 7 AA.  
 ID ALL7\_CYPDPO  
 AC P82158;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Cydlastratin 7.  
 OS Cydia pomonella (Coddling moth).  
 CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 CC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;

OC Ditylidae; Tortricidae; Tortricidae; Olethreutinae; Cydia.  
 OX NCBI\_TaxID=82600;  
 RN (1)  
 RP SEQUENCE.  
 RC TISSUE-Larva;  
 RX MEDLINE=98054539; PubMed=9392829;  
 RA Duve H., Johnson A.H., Maestro J.-L., Scott A.G., Winstanley D.,  
 RT Davey M., East P.D., Thorpe A.;  
 RL "Leptopteran peptides of the allostatin superfamily.";  
 CC Peptides 18:1301-1309(1997).  
 CC -1- SIMILARITY: BELONGS TO THE ALLOSTATIN FAMILY.  
 CC Neuropeptide; Amidation.  
 FT MOD.RES 7  
 SQ SEQUENCE 7 AA; 873 MW; 672879CAB569350 CRC64;  
 Query Match 40.0%; Score 6; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 10 D 10  
 Db 4 D 4  
 RESULT 13  
 CHOX\_ALCSP STANDARD; PRT; 7 AA.  
 ID ID FAR2\_PROCL  
 AC P16101;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 01-APR-1990 (Rel. 14, Last annotation update)  
 DE Choline oxidase (EC 1.1.3.17) (Fragment).  
 OS Alcaligenes sp.  
 OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;  
 OC Alcaligenes  
 NC NCBI\_TaxID=512;  
 RN (1)  
 RP SEQUENCE.  
 RX MEDLINE=81006769; PubMed=6997283;  
 RA Ohta-Fukuyama M., Miyake Y., Eml S., Yamano T.;  
 RT "Identification and properties of the prosthetic group of choline  
 oxidase from Alcaligenes sp.";  
 RL J. Biochem. 88:197-203(1980).  
 CC -1- CATALYTIC ACTIVITY: Choline + O(2) = betaine aldehyde + H(2)O(2).  
 DR PIR: A15398; A15398.  
 KM Oxidoreductase.  
 FT NON\_TER 7  
 SQ SEQUENCE 7 AA; 839 MW; 7415B1E457644AC0 CRC64;  
 Query Match 40.0%; Score 6; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 10 D 10  
 Db 1 D 1  
 RESULT 14  
 FAR1\_HELTI STANDARD; PRT; 7 AA.  
 ID ID FAR1\_HELTI  
 AC P41871;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE FMRFamide-like neuropeptide GDPPLRF-amide.  
 OS Helisoma trivolvis (Snail).  
 CC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;  
 CC Planorbidae; Helisoma.  
 NC NCBI\_TaxID=27815;  
 RN (1)  
 RP SEQUENCE.  
 RC TISSUE-Kidney;

RX MEDLINE=94286417; PubMed=7912428;  
 RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;  
 RT "FMRFamide-related peptides from the kidney of the snail, Helisoma  
 trivolvis.";  
 RL Peptides 15:31-36(1994).  
 CC -1- FUNCTION: APPEARS TO BE INVOLVED IN OSMOREGULATION BY AFFECTING  
 CC THE KIDNEY, MANTLE AND SKIN.  
 CC -1- TISSUE SPECIFICITY: KIDNEY, SKIN, MANTLE AND THE HEMOLYMPH.  
 CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 CC Neuropeptide; Amidation.  
 FT MOD.RES 7  
 SQ SEQUENCE 7 AA; 851 MW; 69D40729D76AA810 CRC64;  
 Query Match 40.0%; Score 6; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 10 D 10  
 Db 2 D 2  
 RESULT 15  
 FAR2\_PROCL STANDARD; PRT; 7 AA.  
 ID ID FAR2\_PROCL  
 AC P38498;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE Cardioexcitatory FMRFamide homolog DF2.  
 OS Procambarus clarkii (Red swamp crayfish).  
 CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;  
 CC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;  
 OC Astacidae; Astacoidea; Cambaridae; Procambarus.  
 NC NCBI\_TaxID=6728;  
 RN (1)  
 RP SEQUENCE.  
 RC TISSUE-Pericardial organs;  
 RX MEDLINE=93248032; PubMed=8387183;  
 RA Mercier A.J., Orchard I., Tedruggue V., Skerrett M.;  
 RT "Isolation of two FMRFamide-related peptides from crayfish  
 pericardial organs.";  
 RL Peptides 14:137-143(1993).  
 CC -1- FUNCTION: INCREASES THE RATE AND AMPLITUDE OF SPONTANEOUS  
 CC CONTRACTIONS OF SEMI-ISOLATED HEARTS. INCREASES THE AMPLITUDE OF  
 CC EXCITATORY POSTSYNAPTIC POTENTIALS IN ABDOMINAL EXTENSOR MUSCLE.  
 CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 CC Neuropeptide; Amidation.  
 FT MOD.RES 7  
 SQ SEQUENCE 7 AA; 967 MW; 69D40729C4540AC0 CRC64;  
 Query Match 40.0%; Score 6; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 10 D 10  
 Db 1 D 1  
 Search completed: June 5, 2003, 15:56:34  
 Job time : 7.51515 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:32:58 ; Search time 29.0909 Seconds  
(without alignments)  
70.829 Million cell updates/sec

Title: US-09-150-947f-15  
Perfect score: 15  
Sequence: 1 xxxxxxxxxd 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp Vertebrate:\*
- 14: sp Unclassified:\*
- 15: sp\_rv:\*
- 16: sp\_bacteriophage:\*
- 17: sp\_archaeal:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	40.0	5	2	P83073 bacillus ce
2	6	40.0	7	2	O07354 synechococ
3	6	40.0	7	5	P83274 macrobrach
4	6	40.0	7	6	Q28742 otycolagus
5	6	40.0	7	10	O49223 glycine max
6	6	40.0	7	11	O63480 rattus norv
7	6	40.0	7	11	O55184 rattus norv
8	6	40.0	8	2	O921E9 neisseria m
9	6	40.0	8	2	O85406 coxiella bu
10	6	40.0	8	2	O9RQ57 buchnera ap
11	6	40.0	8	2	O9RQ49 buchnera ap
12	6	40.0	8	2	O9R772 escherichia
13	6	40.0	8	2	P72279 rhodococcus
14	6	40.0	8	2	O56759 xanthobacte
15	6	40.0	8	2	P83158 anabaena sp
16	6	40.0	8	2	O51594 escherichia

17	6	40.0	8	2	O935R0	O935R0 staphylococ
18	6	40.0	8	2	P83152	P83152 anabaena sp
19	6	40.0	8	2	O9R3X0	O9R3X0 planktothrix
20	6	40.0	8	2	O53790	O53790 streptococ
21	6	40.0	8	2	O8RSR3	O8RSR3 lactobacill
22	6	40.0	8	3	O9HDS4	O9HDS4 aspergillus
23	6	40.0	8	3	O9URB9	O9URB9 saccharomyc
24	6	40.0	8	4	O9P285	O9P285 homo sapien
25	6	40.0	8	4	O9U500	O9U500 homo sapien
26	6	40.0	8	4	O9HAD3	O9HAD3 homo sapien
27	6	40.0	8	4	O15895	O15895 homo sapien
28	6	40.0	8	4	O15900	O15900 homo sapien
29	6	40.0	8	4	O15902	O15902 homo sapien
30	6	40.0	8	4	O9UMC7	O9UMC7 homo sapien
31	6	40.0	8	4	O9UMH9	O9UMH9 homo sapien
32	6	40.0	8	4	O9UD24	O9UD24 homo sapien
33	6	40.0	8	4	O9PDK3	O9PDK3 homo sapien
34	6	40.0	8	4	O9BYR5	O9BYR5 homo sapien
35	6	40.0	8	5	O9NGM5	O9NGM5 toxoplasma
36	6	40.0	8	5	O9UB13	O9UB13 albinaria h
37	6	40.0	8	5	O94695	O94695 physarum po
38	6	40.0	8	5	P83195	P83195 perkinsus a
39	6	40.0	8	5	P83275	P83275 macrobrachl
40	6	40.0	8	5	P83316	P83316 penaeus mon
41	6	40.0	8	5	O9TWH6	O9TWH6 perlinereis
42	6	40.0	8	5	P82686	P82686 periplaneta
43	6	40.0	8	5	P82687	P82687 periplaneta
44	6	40.0	8	6	O9T778	O9T778 canis faml
45	6	40.0	8	6	O8WMS1	O8WMS1 bos taurus

#### ALIGNMENTS

##### RESULT 1

P83073 PRELIMINARY; PRT; 5 AA.

AC P83073;  
 DT 01-OCT-2001 (TREMUREL. 18, Created)  
 DT 01-OCT-2001 (TREMUREL. 18, Last sequence update)  
 DT 01-OCT-2001 (TREMUREL. 18, Last annotation update)  
 DE 88 kDa protein (Fragment).  
 OS Bacillus cereus.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
 OC Bacillaceae; Bacillus.  
 RX NCBI\_TaxID=1396.  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=NCIMB 11796;  
 RA Browne N., Dowds B.C.A.;  
 RL Submitted (JUL-2001) to the SWISS-PROT data bank.  
 FT NON\_TER  
 SQ SEQUENCE 5 AA; 623 MW; 6B01AA336F00000 CRC64;

Query Match 40.0%; Score 6; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
 Db 3 D 3

##### RESULT 2

O07354 PRELIMINARY; PRT; 7 AA.  
 AC O07354;  
 DT 01-JUL-1997 (TREMUREL. 04, Created)  
 DT 01-JUL-1997 (TREMUREL. 04, Last sequence update)  
 DT 01-DEC-2001 (TREMUREL. 19, Last annotation update)  
 DE Nifk (Fragment).  
 GN Nifk.  
 OS Synechococcus sp. (strain PCC 8801 / RF-1) (Cyanothecae PCC 8801).

OC Bacteria; Cyanobacteria; Chroococcales; Cyanothecae.  
 OX NCBI\_TaxID=41431;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-RF-1;  
 RX MEDLINE=99231861; PubMed=10217509;  
 RA Huang T.C., Lin R.F., Chu M.K., Chen H.M.;  
 RT "Organization and expression of nitrogen-fixation genes in the aerobic  
 RT nitrogen-fixing unicellular cyanobacterium *Synechococcus* sp. strain  
 RT RF-1.";  
 RL Microbiology 145:743-753(1999).  
 DR EMBL; AF003700; AAC35193.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 849 MW; 7412C72AA9D5B030 CRC64;  
 Query Match 40.0%; Score 6; DB 2; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
 Db 4 D 4  
 RESULT 3  
 P83274 PRELIMINARY; PRT; 7 AA.  
 AC P83274;  
 DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE FMRamide-like neuropeptide F1P1 (DRNFLRF-amide).  
 OS Macrobrachium rosenbergii (Giant fresh water prawn).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Caridea;  
 OC Palaemonidae; Palaemonidae; Macrobrachium.  
 OX NCBI\_TaxID=79674;  
 RN [1]  
 RP SEQUENCE AND MASS SPECTROMETRY.  
 RC TISSUE-EYESTALK;  
 RA Sitthigorngul P., Sarathongkum W., Jaldechoey S., Longyant S.,  
 RA Sitthigorngul W.;  
 RT "Novel FMRamide-like neuropeptides from the eyestalk of the giant  
 RT freshwater prawn *Macrobrachium rosenbergii*.";  
 RL Comp. Biochem. Physiol. 120B:587-595(1998).  
 CC -1- MASS SPECTROMETRY: MW=965.7; METHOD=MALDI.  
 CC -1- SIMILARITY: BELONGS TO THE FARP (FMRAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 7  
 SQ SEQUENCE 7 AA; 967 MW; 69D40729C4540AC0 CRC64;  
 Query Match 40.0%; Score 6; DB 5; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
 Db 1 D 1  
 RESULT 4  
 Q28742 PRELIMINARY; PRT; 7 AA.  
 AC Q28742;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Alpha-myosin heavy chain (Fragment).  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=84221901; PubMed=6328491;  
 RA Friedman D.J., Umeda P.K., Sinha A.M., Hsu H.-J., Jokoycic S.,  
 RA Rabinowitz M.;  
 RT "Characterization of genomic clones specifying rabbit alpha- and beta-  
 RT ventricular myosin heavy chains.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 81:3044-3048(1984).  
 DR EMBL; K01698; AAA31415.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 916 MW; 6B1B1AA1E69326B0 CRC64;  
 Query Match 40.0%; Score 6; DB 6; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
 Db 5 D 5  
 RESULT 5  
 O49223 PRELIMINARY; PRT; 7 AA.  
 AC O49223;  
 DT 01-JUN-1998 (TREMBlrel. 06, Created)  
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE HMG-1-like protein (Fragment).  
 OS Glycine max (Soybean).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
 OX NCBI\_TaxID=3847;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. ESSEX; TISSUE=ROOT;  
 RX MEDLINE=91367679; PubMed=1891369;  
 RA Laux T., Goldberg R.B.;  
 RT "A plant DNA binding protein shares highly conserved sequence motifs  
 RT with HMG-box proteins.";  
 RL Nucleic Acids Res. 19:4769-4769(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. ESSEX; TISSUE=ROOT;  
 RA Mahalingam R., Knapp H.T.;  
 RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AF047050; AAC03556.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 850 MW; 6AAAAAB378637810 CRC64;  
 Query Match 40.0%; Score 6; DB 10; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
 Db 5 D 5  
 RESULT 6  
 O63480 PRELIMINARY; PRT; 7 AA.  
 AC O63480;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE TR4-NS orphan receptor (Fragment).  
 GN TR4.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-96198747; PubMed-8612486;
RA Yoshikawa T., Makino S., Gao X.M., Xing G.Q., Chuang D.M.,
RT Detera-Wadleigh S.D.;
RT "Splice variants of rat TR4 orphan receptor: differential expression
RT of novel sequences in the 5'-untranslated region and C-terminal
RT domain.";
RL Endocrinology 137:1562-1571(1996).
DR EMBL; U59125; AAB02827.1; -.
FT Receptor.
KW Receptor.
SQ SEQUENCE 7 AA; 758 MW; 672AA87864005350 CRC64;

OY Query Match 40.0%; Score 6; DB 11; Length 7;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10
DB 6 D 6

RESULT 7
O55184 PRELIMINARY; PRT; 7 AA.
ID O55184;
AC O55184;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Orphan receptor TR4-NS (Fragment).
CN TR4.
OC Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RP SEQUENCE FROM N.A.
RX STRAIN-SPRAGUE-DAWLEY;
RA MEDLINE-96198747; PubMed-8612486;
RA Yoshikawa T., Makino S., Gao X.M., Xing G.Q., Chuang D.M.,
RA Detera-Wadleigh S.D.;
RT "Splice variants of rat TR4 orphan receptor: differential expression
RT of novel sequences in the 5'-untranslated region and C-terminal
RT domain.";
RL Endocrinology 137:1562-1571(1996).
[2]
RP SEQUENCE FROM N.A.
RX STRAIN-SPRAGUE-DAWLEY;
RA MEDLINE-96299786; PubMed-8661150;
RA Yoshikawa T., DuPont B.R., Leach R.J., Detera-Wadleigh S.D.;
RT "New variants of the human and rat nuclear hormone receptor, TR4:
RT expression and chromosomal localization of the human gene.";
RL Genomics 35:361-366(1996).
DR EMBL; U59454; AAB91433.1; -.
FT Receptor.
KW Receptor.
SQ SEQUENCE 7 AA; 663 MW; 6DDAA8787EB05350 CRC64;

OY Query Match 40.0%; Score 6; DB 11; Length 7;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10
DB 6 D 6

RESULT 8
O921E9 PRELIMINARY; PRT; 8 AA.
ID O921E9
AC O921E9;
DT 01-MAY-1999 (TREMBLrel. 10, Created)

```

ID	Q9R057	PRELIMINARY;	PRT;	8 AA.
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)			
DT	01-MAY-1999 (TREMBLrel. 10, Last annotation update)			
DE	Cardamoyl-phosphate synthase subunit B (fragment).			
GN	CAB.			
OS	Neisseria meningitidis.			
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.			
OX	NCBI_TaxID=487;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-1527;			
RA	Brinkman F.S.L., Francis F.M., Dillon J.R.;			
RT	"Complexity of the variable sequence between the cardamoyl-phosphate synthase genes of Neisseria species.";			
RT	Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.			
RL	EMBL: AF029361; AAC78449.1; -			
FT	NON_TER			
SO	SEQUENCE 8 AA; 988 MW; FA372AB184032766 CRC64;			
QY	10 D 10			
Db	6 D 6			
DT	01-NOV-1998 (TREMBLrel. 08, Created)			
DT	01-NOV-1998 (TREMBLrel. 08, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	Hypothetical 1.0 kDa protein (Fragment).			
OS	Coxiella burnetii.			
OC	Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;			
OC	Coxiella group; Coxiella.			
OX	NCBI_TaxID=777;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-NINE MILE PHASE I;			
RX	MEDLINE=98348442; PubMed=9683477;			
RA	Willems H., Jaeger C., Baljer G.;			
RT	"Physical and genetic map of the obligate intracellular bacterium Coxiella burnetii.";			
RT	J. Bacteriol. 180:3816-3822(1998).			
RL	EMBL: AF064963; AAD09947.1; -			
KW	Hypothetical protein.			
FT	NON_TER			
SO	SEQUENCE 8 AA; 993 MW; 04GB5AA453772727 CRC64;			
QY	10 D 10			
Db	6 D 6			
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)			
DT	01-MAY-1999 (TREMBLrel. 10, Last annotation update)			
DE	Cardamoyl-phosphate synthase subunit B (fragment).			
GN	CAB.			
OS	Neisseria meningitidis.			
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.			
OX	NCBI_TaxID=487;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-1527;			
RA	Brinkman F.S.L., Francis F.M., Dillon J.R.;			
RT	"Complexity of the variable sequence between the cardamoyl-phosphate synthase genes of Neisseria species.";			
RT	Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.			
RL	EMBL: AF029361; AAC78449.1; -			
FT	NON_TER			
SO	SEQUENCE 8 AA; 988 MW; FA372AB184032766 CRC64;			
QY	10 D 10			
Db	6 D 6			
DT	01-NOV-1998 (TREMBLrel. 08, Created)			
DT	01-NOV-1998 (TREMBLrel. 08, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	Hypothetical 1.0 kDa protein (Fragment).			
OS	Coxiella burnetii.			
OC	Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;			
OC	Coxiella group; Coxiella.			
OX	NCBI_TaxID=777;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-NINE MILE PHASE I;			
RX	MEDLINE=98348442; PubMed=9683477;			
RA	Willems H., Jaeger C., Baljer G.;			
RT	"Physical and genetic map of the obligate intracellular bacterium Coxiella burnetii.";			
RT	J. Bacteriol. 180:3816-3822(1998).			
RL	EMBL: AF064963; AAD09947.1; -			
KW	Hypothetical protein.			
FT	NON_TER			
SO	SEQUENCE 8 AA; 993 MW; 04GB5AA453772727 CRC64;			
QY	10 D 10			
Db	6 D 6			
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)			
DT	01-MAY-1999 (TREMBLrel. 10, Last annotation update)			
DE	Cardamoyl-phosphate synthase subunit B (fragment).			
GN	CAB.			
OS	Neisseria meningitidis.			
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.			
OX	NCBI_TaxID=487;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-1527;			
RA	Brinkman F.S.L., Francis F.M., Dillon J.R.;			
RT	"Complexity of the variable sequence between the cardamoyl-phosphate synthase genes of Neisseria species.";			
RT	Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.			
RL	EMBL: AF029361; AAC78449.1; -			
FT	NON_TER			
SO	SEQUENCE 8 AA; 988 MW; FA372AB184032766 CRC64;			
QY	10 D 10			
Db	6 D 6			
DT	01-NOV-1998 (			

AC 09R057;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE Nifs protein homolog (Fragment).  
 GN NIFS.  
 OS Buchnera aphidicola.  
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.  
 OX NCBI\_TaxID=9;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20022990; PubMed=10555290;  
 RA Clark M.A., Moran N.A., Baumann P.;  
 RT "Sequence evolution in bacterial endosymbionts having extreme base  
 compositions.";  
 RL Mol. Biol. Evol. 16:1586-1598(1999).  
 DR EMBL; AF130812; AAF13797.1; -.  
 FT NON\_TER  
 SQ SEQUENCE 8 AA; 980 MW; F3A73B504771A336 CRC64;

Query Match 40.0%; Score 6; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
 DB 8 D 8

## RESULT 11

09R049 PRELIMINARY; PRT; 8 AA.  
 ID 09R049;  
 AC 09R049;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
 DE Nifs protein homolog (Fragment).  
 GN NIFS.  
 OS Buchnera aphidicola.  
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.  
 OX NCBI\_TaxID=9;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20022990; PubMed=10555290;  
 RA Clark M.A., Moran N.A., Baumann P.;  
 RT "Sequence evolution in bacterial endosymbionts having extreme base  
 compositions.";  
 RL Mol. Biol. Evol. 16:1586-1598(1999).  
 DR EMBL; AF130814; AAF13805.1; -.  
 FT NON\_TER  
 SQ SEQUENCE 8 AA; 992 MW; F3A73B5047672336 CRC64;

Query Match 40.0%; Score 6; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
 DB 8 D 8

## RESULT 12

09R772 PRELIMINARY; PRT; 8 AA.  
 ID 09R772;  
 AC 09R772;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Hypothetical 1.0 kDa protein (Fragment).  
 GN YQFG.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.

OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE=97061202; PubMed=8905232;  
 RA Oshima T., Alda H., Baba T., Fujita K., Hayashi K., Honjo A.,  
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,  
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,  
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,  
 RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,  
 RA Yano M., Horuchi T.;  
 RT "A 718-kb DNA Sequence of Escherichia coli K-12 Genome Corresponding  
 to the 12.7-28.0 min Region on the Linkage Map.";  
 RL DNA Res. 3:137-155(1996).  
 DR EMBL; D90705; BAA5310.1; -.  
 KW Hypothetical protein.  
 FT NON\_TER  
 SQ SEQUENCE 8 AA; 964 MW; DF133B1DD04B476A CRC64;

Query Match 40.0%; Score 6; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
 DB 1 D 1

## RESULT 13

P72279 PRELIMINARY; PRT; 8 AA.  
 ID P72279;  
 AC P72279;  
 DT 01-FEB-1997 (TREMBlrel. 02, Created)  
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Biphenyl dioxygenase (Fragment).  
 GN BPHB.  
 OS Rhodococcus globerulus.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacteriaceae; Nocardiaceae; Rhodococcus.  
 OX NCBI\_TaxID=33008;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95255652; PubMed=7737502;  
 RA Asturias J.A., Diaz E., Timmis K.N.;  
 RT "Evolutionary relationship of the biphenyl dioxygenase of the gram-  
 positive bacterium Rhodococcus globerulus p6 to multicomponent  
 dioxygenases of gram-negative bacteria.";  
 RL Gene 156:11-18(1995).  
 DR EMBL; X80041; CAA56350.1; -.  
 KW Dioxygenase.  
 FT NON\_TER  
 SQ SEQUENCE 8 AA; 989 MW; EBD2CB1AB6D73406 CRC64;

Query Match 40.0%; Score 6; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
 DB 5 D 5

## RESULT 14

056759 PRELIMINARY; PRT; 8 AA.  
 ID 056759;  
 AC 056759;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Halooacid dehalogenase (Fragment).  
 GN DHB.  
 OS Xanthobacter autotrophicus.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Hypnomicrobium group; Xanthobacter.  
 OX NCBI\_TaxID=280;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GJ10, AND CV. M50;  
 RX MEDLINE-95173113; PubMed-7868610;  
 RA Van der Ploeg J., Willemse M., van Hall G., Janssen D.B.;  
 RT "Adaptation of Xanthobacter autotrophicus GJ10 to bromoacetate due to  
 RT activation and mobilization of the haloacetate dehalogenase gene by  
 RT insertion element IS1247";  
 RL J. Bacteriol. 177:1348-1356(1995).  
 DR EMBL; X84038; CAA5857.1; -.  
 FT NON\_TER  
 SQ SEQUENCE 8 AA; 922 MW; F3A9D2D2CDD3056 CRC64;

Query Match 40.0%; Score 6; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
 |  
 Db 8 D 8

## RESULT 15

P83158 PRELIMINARY; PRT; 8 AA.  
 ID P83158;  
 AC P83158;  
 DT 01-DEC-2001 (TRENBLREL. 19, Created)  
 DT 01-DEC-2001 (TRENBLREL. 19, last sequence update)  
 DT 01-DEC-2001 (TRENBLREL. 19, last annotation update)  
 DE Photosystem I Iron-sulfur center (Photosystem I subunit VII) (9 kDa  
 DE polypeptide) (PSI-C) (Fragment).  
 OS Anabaena sp. (strain L31).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.  
 OX NCBI\_TaxID=29412;  
 RN [1]  
 RP SEQUENCE.  
 RA Apte S.K., Uhlemann E., Schmid R., Alendord K.;  
 RL Submitted (OCT-2001) to the SWISS-PROT data bank.  
 CC -1- FUNCTION: APOPROTEIN FOR THE IRON-SULFUR CENTERS FA AND FB OF THE  
 CC PHOTOSYSTEM I COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF  
 CC 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.  
 DR InterPro: IPR001450; 4Fe4S\_Ferredoxin.  
 DR PROSITE: PS00198; 4Fe4S\_FERREDOXIN; PARTIAL.  
 KW Photosynthesis; Photosystem I; Iron-sulfur; 4Fe-4S; Metal-binding.  
 FT NON\_TER  
 SQ SEQUENCE 8 AA; 962 MW; C5B850532D1A1F5 CRC64;

Query Match 40.0%; Score 6; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
 |  
 Db 8 D 8

Search completed: June 5, 2003, 15:55:44  
 Job time : 30.0909 secs

**THIS PAGE BLANK (USPTO)**



XX Claim 2; Fig 2; 72pp; English.

PS

CC The present sequence represents the SMEZ-2 superantigen protein. The

CC specification describes superantigen proteins SMEZ-2, SPE-G, SPE-H

CC and SPE-J. The superantigen polynucleotides and polypeptides are

CC used for subtyping Streptococci. They are also used for diagnosing

CC Streptococcal disease. The superantigens are used in diagnosis of

CC disease such as Kawasaki syndrome. They are also useful to recruit

CC and activate T cells in a relatively non-specific fashion since

CC they bind a large number of T cell receptor molecules by binding to the

CC Vbeta domain. Superantigen constructs are useful in cancer therapy.

XX

SQ Sequence 233 AA;

Query Match 72.7%; Score 24; DB 21; Length 233;

Best Local Similarity 50.0%; Pred. No. 73;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12

DB 149 KTTVTAQED 158

RESULT 2

ABP29565

ID ABP29565 standard; Protein; 258 AA.

XX

AC ABP29565;

XX

DT 02-JUL-2002 (first entry)

XX

DE Streptococcus polypeptide SEQ ID NO 8306.

XX

DE Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;

KM group A streptococcus; Streptococcus pyogenes; antibacterial;

KM antiinflammatory; Infection; vaccine; meningitis; gene therapy.

XX

OS Streptococcus pyogenes.

XX

PN WO200234771-A2.

XX

PD 02-MAY-2002.

XX

PF 29-OCT-2001; 2001WO-GB04789.

XX

PR 27-OCT-2000; 2000GB-0026333.

PR 24-NOV-2000; 2000GB-0028727.

PR 07-MAR-2001; 2001GB-0005640.

XX

PA (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

XX

PI Telford J, Maignant V, Margalit Ros YI, Grandt G, Fraser C;

PI Teteltn H;

XX

DR WPI: 2002-352536/38.

DR N-PSDB; ABN70196.

XX

PT New Streptococcus protein for the treatment or prevention of infection

PT or disease caused by Streptococcus bacteria, such as meningitis, and

PT for detecting a compound that binds to the protein -

XX

PS Claim 1; Page 3947; 4525pp; English.

XX

CC The invention relates to a protein (ABP25413-ABP30895) from group B

CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS

CC (Streptococcus pyogenes), comprising one of 5463 sequences (S1), given in

CC the specification. The proteins have antibacterial and antiinflammatory

CC activity. (1), nucleic acids encoding (1), ABN6044-ABN71526 and

CC antibodies that bind (1) are used in the manufacture of medicaments for

CC the treatment or prevention of infection or disease caused by

CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

CC Nucleic acids encoding (1) are used to detect Streptococcus in a

CC biological sample. (1) is used to determine whether a compound binds to

CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be

CC used as a vaccine or diagnostic composition. The disease caused by

CC Streptococcus that is prevented or treated may be meningitis. Nucleic

CC acid encoding (1) may be used to recombinantly produce (1) and may be

CC used in gene therapy. Antibodies to (1) are used for affinity

CC chromatography, immunoassays, and distinguishing/identifying

CC Streptococcus proteins.

XX

SQ Sequence 258 AA;

Query Match 72.7%; Score 24; DB 23; Length 258;

Best Local Similarity 50.0%; Pred. No. 81;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12

DB 174 KTTVTAQED 183

RESULT 3

ABB76254

ID ABB76254 standard; Peptide; 17 AA.

XX

AC ABB76254;

XX

DT 09-AUG-2002 (first entry)

XX

DE Staphylococcus aureus enterotoxin C1 peptide fragment.

XX

KM Enterotoxin C; SECI; superantigen; antigen; tumour; cancer;

KM antitumour; therapy.

XX

OS Staphylococcus aureus.

XX

PN US2002051765-A1.

XX

PD 02-MAY-2002.

XX

PF 19-DEC-2000; 2000US-0741503.

XX

PR 31-JAN-1994; 94US-0189424.

PR 19-JUN-1995; 95US-0491746.

PR 03-OCT-1989; 89US-0416530.

PR 17-JAN-1990; 90US-0466577.

PR 17-JAN-1991; 91WO-US00342.

PR 01-JUN-1992; 92US-0891718.

PR 02-MAR-1993; 93US-0025144.

XX

PA (TERM/) TERMAN D S.

XX

PI Terman DS;

XX

DR WPI: 2002-415198/44.

XX

PT Reagent for treating cancer without the need for e.g. radiotherapy,

PT comprises a specific V beta subset of T cells sensitized to a growing

PT tumor and stimulated with superantigens -

XX

PS Disclosure; Page 4; 17pp; English.

XX

CC The present sequence corresponds to amino acids 151-167 of

CC enterotoxin C1 (SECI, see ABB76238) of Staphylococcus aureus.

CC Sequence similarities are found in the corresponding regions

CC of other staphylococcal enterotoxins and streptococcal pyrogenic

CC exotoxins (see ABB76252-59). In the present invention, synthetic

CC polypeptides useful in tumour therapy and in blocking or destroying

CC autoreactive T and B lymphocyte populations are characterised by

CC substantial structural homology to staphylococcal enterotoxin A and

CC enterotoxin B, and to streptococcal pyrogenic exotoxins, with

CC statistically significant sequence homology and similarity,

CC including alignment of cysteine residues and similar hydropathy



CC profiles. These superantigens are used to treat solid tumours, including their metastases, without radiation, surgery or standard chemotherapeutic agents. A claimed method of human cancer treatment involves contacting haematopoietic cells from a patient with one or more superantigens *ex vivo* to generate stimulated cells, selecting a specific V beta subset of cells, and reintroducing these cells into the patient to induce an *in vivo* therapeutic, tumouricidal reaction.

XX .Sequence 17 AA;

XX SQ

XX Query Match 69.7%; Score 23; DB 23; Length 17;

XX Best Local Similarity 50.0%; Pred. No. 9.8;

XX Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0

YY 3 KXXXTXQEXD 12

YY | | | |

YY 1 KKSVAQELD 10

DB

RESULT 4

AAAY97855

AAAY97855 standard; peptide: 24 AA.

AC AAY97855;

XX 29-AUG-2000 (first entry)

XX Staphylococcus aureus enterotoxin SEC, conserved region 2.

DE Staphylococcus aureus enterotoxin; septic shock; toxic shock;

KW vaccine; antibody.

OS Staphylococcus aureus.

PN W0200020598-A1.

PD 13-APR-2000.

PF 24-SEP-1999; 99WO-US22180.

PR 07-OCT-1998; 98US-0168303.

PR 18-JUN-1999; 99US-033581.

PA (UYRQ ) UNIV ROCKEFELLER.

PI Bannan JD, Visvanathan K, Zabriskie JB;

DR WPI: 2000-303782/26.

XX Peptides useful for preventing and reducing the symptoms of toxic shock syndrome and septic shock from staphylococcal and streptococcal infections -

XX Example 1: Page 72; 115pp; English.

XX The invention relates to novel peptides (AAY97838-Y97843) comprising a consensus amino acid sequence derived from two conserved regions (regions 1 and 2) of Staphylococcus aureus enterotoxins and streptococcal pyrogenic toxins. Consensus region 1a (a preferred consensus region 1) has the sequence X25-X26-Y-G-G-X1-T-X2-X3-X4-X5-N (AAY97863) and consensus region 2a (a preferred consensus region 2) has the sequence K-X6-X7-X8-X9-X10-X11-X12-X13-D-X14-X15-X16-R-X17-X18-X27-X19-X20-X21-X22-X23-X24-Y (AAY97864), where: X1, X8, X13 and X24 are each independently selected from L, I and V; X2, X4, X5, X6, X7, X9, X10, X11, X12, X14, X15, X16, X17, X18, X19, X20, X21, X22 and X23 are each independently selected from the group consisting of any amino acid; X3, X5 and X26 are each independently selected from the group consisting of any amino acid and of no amino acid; X27 is either L or Y.

XX The invention also relates to serum antibodies induced by the peptides which provide protection against, or reduce the severity of toxic shock

Query Match	Best Local Similarity	Score 23;	DB 21;	Length 24;
Matches	5; Conservative	0; Mismatches	5; Indels	0; Gaps
0Y	3 KXXXTXQEXD 12			
	1 1 1 1			
Db	1 KKSVTQAEILD 10			
Sequence	24 AA:			
Query Match	69.7%; Score 23; <td>DB 21;<td>Length 24;<td></td></td></td>	DB 21; <td>Length 24;<td></td></td>	Length 24; <td></td>	
Best Local Similarity	50.0%; Pred. No. 14;			
Matches	5; Conservative	0; Mismatches	5; Indels	0; Gaps
RESULT 5				
AAV72193				
ID	AAV72193 standard; peptide; 24 AA.			
AAV72193;				
24-APR-2001	(first entry)			
Peptide from region 2 of staphylococcal enterotoxin SEC.				
Mononuclear cell; blastogenesis; inhibitor; HIV; replication; therapy;				
staphylococcal enterotoxin; human immunodeficiency virus; T cell;				
autoimmune disease; immunisation.				
Staphylococcus sp.				
MO200078790-A2.				
28-DEC-2000.				
16-JUN-2000; 2000WO-US16680.				
18-JUN-1999; 99US-0336627.				
(UTRQ ) UNIV ROCKEFELLER.				
Visvanathan K, Zabriskie JB;				
WPI: 2001-080820/09.				
Providing protection against, and reducing the severity of, human				

PT immunodeficiency virus infections and associated deleterious effects,  
PT using peptides from homologous sequences of staphylococcal and  
XX streptococcal toxins  
XX  
PS Disclosure; Page 36; 76pp; English.  
XX  
CC The present sequence is a peptide from region 2 of staphylococcal  
CC enterotoxin SEC.  
CC The peptide, nucleic acid encoding the peptide and antibody (Ab) produced  
CC against the peptide are useful for inhibiting biogenesis of  
CC mononuclear cells (eg. T cells) in the presence of human immunodeficiency  
CC virus (HIV), inhibiting HIV replication and protecting a mammal against  
CC the deleterious effects of HIV. The peptide is also used to ameliorate  
CC the effects of autoimmune diseases associated with the presence of HIV.  
CC The Ab is used for passively immunising a mammal against the deleterious  
CC effects of HIV.  
XX  
SQ Sequence 24 AA:  
  
Query Match 69.7%; Score 23; DB 22; Length 24;  
Best Local Similarity 50.0%; Pred. No. 14;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
  
OY 3 KXXXTXQEXD 12  
| | | | |  
DB 1 KKSVAQDEL 10  
  
RESULT 6  
AAM59147  
ID AAM59147 standard; Protein; 29 AA.  
XX  
AC AAM59147;  
XX  
DT 05-NOV-2001 (first entry)  
XX  
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 31252.  
XX  
KW Human; brain expressed; exon; gene expression analysis; probe;  
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
KW epilepsy; cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO200157275-A2.  
XX  
PD 09-AUG-2001.  
XX  
PE 30-JAN-2001; 2001WO-US00667.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-483446/52.  
XX  
PT Single exon nucleic acid probes for analyzing gene expression in human  
PT brains -  
XX  
PS Example 4; SEQ ID NO: 31252; 650pp + Sequence Listing; English.  
XX  
CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is a protein encoded by one of  
CC the probes of the invention.  
XX  
SQ Sequence 29 AA:  
  
Query Match 69.7%; Score 23; DB 22; Length 29;  
Best Local Similarity 50.0%; Pred. No. 17;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
  
OY 3 KXXXTXQEXD 12  
| | | | |  
DB 10 KALSTGDEL 19  
  
RESULT 7  
AAM71683  
ID AAM71683 standard; Protein; 29 AA.  
XX  
AC AAM71683;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 31989.  
XX  
KW Human; bone marrow expressed exon; gene expression analysis; probe;  
KW microarray; cancer; leukemia; lymphoma; myeloma.  
XX  
OS Homo sapiens.  
XX  
PN WO200157276-A2.  
XX  
PD 09-AUG-2001.  
XX  
PE 30-JAN-2001; 2001WO-US00668.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488900/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human bone marrow -  
XX  
PS Example 4; SEQ ID NO: 31989; 658pp + Sequence Listing; English.  
XX  
CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC bone marrow. They can be used to measure gene expression in bone marrow  
CC samples, which may enable the improved diagnosis and treatment of cancers  
CC such as lymphoma, leukemia and myeloma. The present sequence is a  
CC protein encoded by one of the probes of the invention.  
XX  
SQ Sequence 29 AA:  
  
Query Match 69.7%; Score 23; DB 22; Length 29;  
Best Local Similarity 50.0%; Pred. No. 17;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
  
OY 3 KXXXTXQEXD 12  
| | | | |  
DB 10 KALSTGDEL 19

RESULT 8  
 AAM31976  
 ID AAM31976 standard; Protein: 29 AA.  
 XX  
 AC AAM31976;  
 XX  
 DT 17-OCT-2001 (first entry)  
 XX  
 DE Peptide #6013 encoded by probe for measuring placental gene expression.  
 XX  
 KW Probe; microarray; human; placenta; antenatal diagnosis;  
 KW genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157272-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00663.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI: 2001-488897/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human placenta -  
 XX  
 PS Claim 27; SEQ ID NO 32245; 654pp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes (SENP:  
 CC see A131315-A157546). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for producing a microarray for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from human placenta. The probes are useful for antenatal diagnosis of  
 CC human genetic disorders.  
 CC  
 SQ Sequence 29 AA:  
 Query Match 69.7%; Score 23; DB 22; Length 29;  
 Best Local Similarity 50.0%; Pred. No. 17;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 OY 3 KAXXTXQEXD 12  
 DB 10 KALSTGOEID 19  
 RESULT 9  
 ABG41496  
 ID ABG41496 standard; Peptide: 29 AA.  
 XX  
 AC ABG41496;  
 XX  
 DT 19-AUG-2002 (first entry)  
 XX  
 DE Human peptide encoded by genome-derived single exon probe SEQ ID 31161.  
 XX  
 KW Human: single exon probe; asthma; lung cancer; COPD; ILD;  
 KW chronic obstructive pulmonary disease; interstitial lung disease;  
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;

KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;  
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
 KW primary ciliary dyskinesia; pulmonary hypertension;  
 KW hyaline membrane disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200186003-A2.  
 XX  
 PD 15-NOV-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00665.  
 XX  
 PR 04-FEB-2000; 2000US-180312P.  
 PR 26-MAY-2000; 2000US-207456P.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-234687P.  
 PR 27-SEP-2000; 2000US-236359P.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI: 2002-114183/15.  
 XX  
 PT Spatially-addressable set of single exon nucleic acid probes, used to  
 PT measure gene expression in human lung samples -  
 XX  
 PS Claim 27; SEQ ID NO 31161; 634pp; English.  
 XX  
 CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human lung comprising single exon nucleic acid probes having one of  
 CC 12614 nucleic acid sequences mentioned in the specification, or their  
 CC complements or the 12387 open reading frames derived from the 12614  
 CC probes. Also included are a microarray comprising the novel set of  
 CC probes; the novel set of probes which hybridize at high stringency to a  
 CC nucleic acid expressed in the human lung; measuring gene expression to a  
 CC sample derived from human lung, comprising (a) contacting the array with  
 CC a collection of detectably labeled nucleic acids derived from human lung  
 CC mRNA, and (b) measuring the label detectably bound to each probe of  
 CC the array; identifying exons in a eukaryotic genome, comprising  
 CC (a) algorithmically predicting at least one exon from genomic sequences  
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably  
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
 CC having a fragment identical to the predicted exon, the probe is included  
 CC in the above mentioned microarray; assigning exons to a single gene,  
 CC comprising (a) identifying exons from genomic sequence by the method  
 CC above and (b) measuring the expression of each of the exons in several  
 CC tissues and/or cell types using hybridisation to a single exon  
 CC microarrays having a probe with the exon, where a common pattern of  
 CC expression of the exons in the tissues and/or cell types indicates that  
 CC the exons should be assigned to a single gene; a peptide comprising one  
 CC of 12011 sequences, mentioned in the specification, or encoded by the  
 CC probes/open reading frames (ORF). The probes are used for gene  
 CC expression analysis, and for identifying exons in a gene, particularly  
 CC using human lung derived mRNA and for the study of lung diseases  
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease  
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,  
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary  
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,  
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic  
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension  
 CC and hyaline membrane disease. The present sequence is a peptide/protein  
 CC encoded by a single exon probe of the invention.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 XX ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 29 AA;  
 Query Match 69.7%; Score 23; DB 23; Length 29;  
 Best Local Similarity 50.0%; Pred. No. 17;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 3 KXXXTXQEXD 12  
 DB 10 KALSTGOEID 19  
 RESULT 10  
 AAG63856  
 ID AAG63856 standard; Protein: 228 AA.  
 AC AAG63856;  
 DT 29-OCT-2001 (first entry)  
 DE Amino acid sequence of a modified Staphylococcal enterotoxin C1.  
 XX Enterotoxin C1; SEC-SER; Infectious disease; mastitis.  
 XX Synthetic.  
 OS Staphylococcus sp.  
 XX WO200160851-A1.  
 PN 23-AUG-2001.  
 PD 31-OCT-2000; 2000WO-KR01241.  
 PF 17-FEB-2000; 2000KR-0007612.  
 PR (GLDS ) LG CHEM LTD.  
 XX Lee H, Park Y, Han K, Chang B, Lee Y;  
 PI WPI: 2001-522585/57.  
 DR N-PSDB: AAH74983.  
 XX  
 PT Producing stable modified Staphylococcal toxin polypeptide for treating  
 PT infectious diseases, e.g. mastitis, in animals, involves substituting  
 PT cysteine at specified position of modified Staphylococcal toxin C1 with  
 PT serine -  
 PS Claim 1; Page 59-60; 64pp; English.  
 XX  
 CC The present sequence represents a modified Staphylococcal enterotoxin C1,  
 CC designated SEC-SER. The modified toxin is characterised in that the  
 CC 95th amino acid (cysteine) is substituted with serine. This results in  
 CC a toxin that has improved stability. The modified enterotoxin is  
 CC useful for preventing or treating infectious diseases due to  
 CC microorganisms in animals such as cows, pigs, horses, sheep, hens,  
 CC dogs and cats, e.g. mastitis in cows.  
 XX  
 SQ Sequence 228 AA;  
 Query Match 69.7%; Score 23; DB 22; Length 228;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 3 KXXXTXQEXD 12  
 DB 140 KKSVTAGQELD 149  
 RESULT 11  
 AAR13208  
 ID AAR13208 standard; Protein: 238 AA.  
 AC AAR13208;  
 XX

DT 15-OCT-1991 (first entry)  
 DE Staphylococcal enterotoxin C3.  
 XX SEC3; cancer treatment; pyrogen; tumouricide.  
 KW Staphylococcus aureus.  
 OS WO9110680-A.  
 PN 25-JUL-1991.  
 PD 17-JAN-1991; 91WO-US00342.  
 PF 17-JAN-1990; 90US-046577.  
 PR (TERM/) TERMAN D S.  
 PA Terman DS;  
 PI WPI: 1991-237984/32.  
 DR  
 XX  
 PT Treating cancer with enterotoxin from Staphylococcus aureus -  
 PT administered by IV injection, having same tumoricidal activity  
 PT as Staphylococcal protein A without potential toxic reactions  
 XX  
 PS Disclosure: Fig 1; 74pp; English.  
 XX  
 CC SEC3 was isolated and purified from S.aureus. It can be used for  
 CC treating cancer, activating cytokine mediators and procoagulant  
 CC systems, augmenting natural killer cell cytotoxicity, etc. The  
 CC enterotoxin can be administered intravenously, optionally with  
 CC ibuprofen to attenuate toxic reaction to SEC3. Synthetic  
 CC polypeptides having structural homology to Staphylococcal exotoxins  
 CC are claimed, provided the homology includes statistically  
 CC significant sequence homology, alignment of Cysteine residues and  
 CC similar hydrophaty profiles.  
 CC See AAR13203-R13211.  
 XX  
 SQ Sequence 238 AA;  
 Query Match 69.7%; Score 23; DB 12; Length 238;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 3 KXXXTXQEXD 12  
 DB 150 KKSVTAGQELD 159  
 RESULT 12  
 AAR45016  
 ID AAR45016 standard; Protein: 238 AA.  
 AC AAR45016;  
 XX  
 DT 08-JUN-1994 (first entry)  
 DE Staphylococcal enterotoxin SEC3.  
 DE Staphylococcal enterotoxin; SE; cancer; tumouricidal agent;  
 KW autoimmune disease; toxicity; Protein A; perfusion system.  
 XX Staphylococcus aureus.  
 OS WO9324136-A.  
 PN 09-DEC-1993.  
 PD 01-JUN-1993; 93WO-US05213.  
 PF 01-JUN-1992; 92US-0891718.  
 PR  
 XX

PA (STONE/) STONE J L.  
 (TERM/) TERMAN D S.  
 XX Stone JL, Terman DS;  
 PI WPI: 1993-405418/50.  
 DR WPI: 1993-405418/50.  
 XX  
 XX Use of staphylococcal enterotoxin(s) and homologues - for  
 PT treating cancer in a patient or for the treatment of auto-immune  
 PT diseases  
 XX  
 XX Disclosure: Fig 1: 90pp: English.  
 XX  
 CC The sequences given in AAR45011-21 are Staphylococcal enterotoxins (SEs)  
 CC which may be used in the methods for the invention for treating cancer  
 CC in a patient. These SEs, and homologues of them, can be used as  
 CC tumoricidal agents for treating cancers and autoimmune disease.  
 CC They exhibit tumoricidal activity and toxicity identical to that  
 CC observed for the Protein A pertussis system. They may be administered  
 CC by i.v. injection.  
 CC  
 SQ Sequence 238 AA:  
 Query Match 69.7%; Score 23; DB 14; Length 238;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 OY 3 KXXXTXQEXD 12  
 DB 150 KSVTAQELD 159  
 RESULT 13  
 AAB67343  
 ID AAB67343 standard; peptide; 238 AA.  
 AC AAB67343;  
 XX  
 DT 23-APR-2001 (first entry)  
 XX  
 DE Staphylococcus aureus enterotoxin C3 protein.  
 XX  
 KM Tumour; cancer; immune; enterotoxin.  
 XX  
 OS Staphylococcus aureus.  
 XX  
 PN US6180097-B1.  
 PD 30-JAN-2001.  
 XX  
 PF 30-OCT-1998; 98US-0183437.  
 XX  
 PR 31-JAN-1994; 94US-0189424.  
 PR 19-JUN-1995; 95US-0491746.  
 PR 03-OCT-1989; 89US-0416530.  
 PR 17-JAN-1990; 90US-0466577.  
 PR 17-JAN-1991; 91WO-US00342.  
 PR 01-JUN-1992; 92US-0891718.  
 PR 02-MAR-1993; 93US-0025144.  
 XX  
 (TERM/) TERMAN D S.  
 PA  
 XX Terman DS;  
 PI WPI: 2001-158657/16.  
 DR  
 XX  
 XX Tumor cell capable of stimulating antitumor immune reactivity in vitro  
 PT or in vivo comprises exogenous nucleic acids encoding a superantigen  
 PT and a costimulatory molecule -  
 XX  
 PS Disclosure: Fig 2: 16pp: English.  
 CC The present invention relates to a tumour cell capable of stimulating

CC antitumor immune reactivity in vitro or in vivo contains and  
 CC expresses an exogenous nucleic acid molecule encoding a superantigen  
 CC or its active fragment and an exogenous nucleic acid molecule  
 CC encoding a costimulatory molecule that activates T cells in  
 CC conjunction with an antigenic stimulus. The invention may be used  
 CC for cancer therapy by stimulating an anticancer immune response  
 CC in vivo or ex vivo.  
 CC  
 SQ Sequence 238 AA:  
 Query Match 69.7%; Score 23; DB 22; Length 238;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 OY 3 KXXXTXQEXD 12  
 DB 150 KSVTAQELD 159  
 RESULT 14  
 ABB76239  
 ID ABB76239 standard; Protein; 238 AA.  
 AC ABB76239;  
 XX  
 DT 09-AUG-2002 (first entry)  
 XX  
 DE Staphylococcus aureus enterotoxin C3.  
 XX  
 KM Enterotoxin C3; SEC3; superantigen; antigen; tumour; cancer;  
 KM antitumour; therapy.  
 XX  
 OS Staphylococcus aureus.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 208 /note= "given as 'O' in the specification"  
 PN US2002051765-A1.  
 PD 02-MAY-2002.  
 XX  
 PF 19-DEC-2000; 2000US-0741503.  
 XX  
 PR 31-JAN-1994; 94US-0189424.  
 PR 19-JUN-1995; 95US-0491746.  
 PR 03-OCT-1989; 89US-0416530.  
 PR 17-JAN-1990; 90US-0466577.  
 PR 17-JAN-1991; 91WO-US00342.  
 PR 01-JUN-1992; 92US-0891718.  
 PR 02-MAR-1993; 93US-0025144.  
 XX  
 (TERM/) TERMAN D S.  
 PA  
 XX Terman DS;  
 PI WPI: 2002-415198/44.  
 DR  
 XX  
 XX Reagent for treating cancer without the need for e.g. radiotherapy,  
 PT comprises a specific V beta subset of T cells sensitized to a growing  
 PT tumor and stimulated with superantigens -  
 XX  
 PS Disclosure: Fig 2: 17pp: English.  
 CC The present sequence is the protein sequence of enterotoxin C3  
 CC (SEC3) of Staphylococcus aureus. Similarly is shown, in several  
 CC stretches of sequence, between staphylococcal enterotoxins,  
 CC streptococcal, pyrogenic exotoxins and staphylococcal exfoliative  
 CC toxins (see ABB76234-44). In the present invention, synthetic  
 CC polypeptides useful in tumour therapy and in blocking or destroying  
 CC autoreactive T and B lymphocyte populations are characterised by  
 CC substantial structural homology to staphylococcal enterotoxin A and  
 CC enterotoxin B, and to streptococcal pyrogenic exotoxins, with

CC statistically significant sequence homology and similarity (% value  
 CC of Lipman and Pearson algorithm in Monte Carlo analysis exceeding  
 CC 6) to include alignment of cysteine residues and similar hydrophathy  
 CC profiles. These superantigens are used to treat solid tumours,  
 CC including their metastases, without radiation, surgery or  
 CC standard chemotherapeutic agents. A claimed method of human cancer  
 CC treatment involves contacting haematopoietic cells from a patient  
 CC with one or more superantigens ex vivo to generate stimulated cells,  
 CC selecting a specific V beta subset of cells, and reintroducing  
 CC these cells into the patient to induce an in vivo therapeutic,  
 CC tumoricidal reaction.

XX  
 SQ Sequence 238 AA;

Query Match 69.7%; Score 23; DB 23; Length 238;  
 Best Local Similarity 50.0%; Pred. No. 1.4e+02;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12  
 | | | |  
 DB 150 KKSVTQAQELD 159

# RESULT 15

AAR13207 standard; Protein; 239 AA.

XX AAR13207;

DT 15-OCT-1991 (first entry)

DE Staphylococcal enterotoxin C1.

XX SECI; cancer treatment; pyrogen; tumouricide.

OS Staphylococcus aureus.

PN W09110680-A.

PD 25-JUL-1991.

PF 17-JAN-1991; 91WO-US00342.

PR 17-JAN-1990; 90US-0466577.

XX (TERM/) TERMAN D S.

PI Terman DS;

DR WPI; 1991-237984/32.

PT Treating cancer with enterotoxin from Staphylococcus aureus -  
 PT administered by IV injection, having same tumoricidal activity  
 PT as Staphylococcal protein A without potential toxic reactions

PS Disclosure; Fig 1; 74pp; English.

XX SECI was isolated and purified from S.aureus. It can be used for  
 CC treating cancer, activating cytokine mediators and procoagulant  
 CC systems augmenting natural killer cell cytotoxicity, etc. The  
 CC enterotoxin can be administered intravenously, optionally with  
 CC lipothen to attenuate toxic reaction to SECI. Synthetic  
 CC polypeptides having structural homology to SECI. Synthetic  
 CC are claimed, provided the homology includes statistically  
 CC significant sequence homology, alignment of Cysteine residues and  
 CC similar hydrophathy profiles.  
 CC See AAR13203-R13211.

XX SQ Sequence 239 AA;

Query Match 69.7%; Score 23; DB 12; Length 239;

Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12  
 | | | |  
 DB 151 KKSVTQAQELD 160

Search completed: June 5, 2003, 16:00:25  
 Job time : 40.7273 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:23:33 ; Search time 10.5455 seconds  
(without alignments)  
33,481 Million cell updates/sec

Title: US-09-150-947F-16  
Perfect score: 33  
Sequence: 1 XXXXXXTXQEXD 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCITUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	23	69.7	17	4	US-08-896-933-6
2	23	69.7	17	4	US-09-314-235-6
3	23	69.7	24	3	US-08-838-413A-20
4	23	69.7	238	4	US-08-896-933-28
5	23	69.7	238	4	US-09-314-235-28
6	23	69.7	239	4	US-08-896-933-27
7	23	69.7	239	4	US-09-314-235-27
8	23	69.7	266	4	US-09-144-776B-14
9	22	66.7	17	4	US-08-896-933-4
10	22	66.7	17	4	US-08-896-933-12
11	22	66.7	17	4	US-08-896-933-14
12	22	66.7	17	4	US-08-896-933-15
13	22	66.7	17	4	US-08-896-933-16
14	22	66.7	17	4	US-08-896-933-17
15	22	66.7	17	4	US-09-314-235-4
16	22	66.7	17	4	US-09-314-235-12
17	22	66.7	17	4	US-09-314-235-14
18	22	66.7	17	4	US-09-314-235-15
19	22	66.7	17	4	US-09-314-235-16
20	22	66.7	17	4	US-09-314-235-17
21	22	66.7	17	4	US-08-838-413A-19
22	22	66.7	24	3	US-08-838-413A-24
23	22	66.7	191	4	US-08-858-207A-381
24	22	66.7	221	4	US-08-896-933-29
25	22	66.7	221	4	US-09-314-235-29
26	22	66.7	239	4	US-08-896-933-26
27	22	66.7	239	4	US-09-314-235-26

28	22	66.7	239	4	US-09-144-776B-10	Sequence 10, Appl
29	22	66.7	251	4	US-09-144-776B-16	Sequence 16, Appl
30	22	66.7	255	1	US-08-446-918A-2	Sequence 2, Appl
31	22	66.7	255	2	US-08-580-806-2	Sequence 8, Appl
32	22	66.7	266	4	US-09-414-276-8	Sequence 2, Appl
33	22	66.7	266	4	US-09-144-776B-6	Sequence 6, Appl
34	22	66.7	266	4	US-09-144-776B-8	Sequence 8, Appl
35	22	66.7	292	4	US-09-190-824-2	Sequence 2, Appl
36	22	66.7	922	2	US-08-464-402-2	Sequence 2, Appl
37	22	66.7	922	4	US-09-054-775C-2	Sequence 2, Appl
38	21	63.6	17	4	US-08-896-933-2	Sequence 2, Appl
39	21	63.6	17	4	US-08-896-933-8	Sequence 8, Appl
40	21	63.6	17	4	US-08-896-933-10	Sequence 10, Appl
41	21	63.6	17	4	US-09-314-235-2	Sequence 2, Appl
42	21	63.6	17	4	US-09-314-235-8	Sequence 8, Appl
43	21	63.6	17	4	US-09-314-235-10	Sequence 10, Appl
44	21	63.6	24	3	US-08-838-413A-4	Sequence 4, Appl
45	21	63.6	24	3	US-08-838-413A-18	Sequence 18, Appl

## ALIGNMENTS

```

RESULT 1
US-08-896-933-6
; Sequence 6, Application US/08896933
; Patent No. 6221351
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; FILE REFERENCE: 09629/005002
; CURRENT APPLICATION NUMBER: US/08/896,933
; EARLIER APPLICATION NUMBER: 1997-07-18
; EARLIER FILING DATE: 1994-06-02
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-896-933-6
;
;
Query Match      69.7%; Score 23; DB 4; Length 17;
Best Local Similarity 50.0%; Pred. No. 0.86;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY      3 KXXXTXQEXD 12
DB      1 KKSXTAQELD 10
;
RESULT 2
US-09-314-235-6
; Sequence 6, Application US/09314235
; Patent No. 6338845
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; FILE REFERENCE: 09629/005004
; CURRENT APPLICATION NUMBER: US/09/314,235
; EARLIER APPLICATION NUMBER: 1999-05-18
; EARLIER FILING DATE: 1997-07-18
; EARLIER FILING DATE: 1994-06-02
; EARLIER APPLICATION NUMBER: 08/252,978
; EARLIER FILING DATE: 1994-06-02
; EARLIER APPLICATION NUMBER: 07/891,718
; EARLIER FILING DATE: 1992-06-01
; EARLIER APPLICATION NUMBER: US91/00342
; EARLIER FILING DATE: 1991-01-17
; EARLIER APPLICATION NUMBER: 07/466,577

```

EARLIER FILING DATE: 1990-01-17  
EARLIER APPLICATION NUMBER: 07/416,530  
EARLIER FILING DATE: 1989-10-03  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 6  
LENGTH: 17  
TYPE: PRT  
ORGANISM: Staphylococcus aureas  
US-09-314-235-6

Query Match  
Best Local Similarity 69.7%; Score 23; DB 4; Length 17;  
50.0%; Pred. No. 0.86;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
| | | | |  
DB 1 KKSVAQELD 10

RESULT 3  
US-08-838-413A-20  
Sequence 20, Application US/08838413A  
Patent No. 6075119  
GENERAL INFORMATION:

APPLICANT:  
TITLE OF INVENTION: PEPTIDES USEFUL FOR  
REDUCING SYMPTOMS OF TOXIC SHOCK SYNDROME  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & PINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154

COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERECT 5.1  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/838,413A  
FILING DATE: 07-APR-1997

CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:

CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:

NAME: MORRIS, MARY J.  
REGISTRATION NUMBER: 34,398

REFERENCE/DOCKET NUMBER: 2016-4010

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)758-4800

TELEFAX: (212)751-6849

INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:

LENGTH: 24

TYPE: AMINO ACID

STRANDEDNESS: UNKNOWN

TOPOLOGY: UNKNOWN

MOLECULE TYPE: PEPTIDE

US-08-838-413A-20

Query Match  
Best Local Similarity 69.7%; Score 23; DB 3; Length 24;  
50.0%; Pred. No. 1.3;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
| | | | |  
DB 1 KKSVAQELD 10

RESULT 4  
US-08-896-933-28  
Sequence 28, Application US/08896933  
Patent No. 6221351  
GENERAL INFORMATION:

APPLICANT: Terman, David S.

TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,  
AND RELATED COMPOUNDS

FILE REFERENCE: 09629/005002

CURRENT APPLICATION NUMBER: US/08/896,933

EARLIER FILING DATE: 1997-07-18

EARLIER APPLICATION NUMBER: 08/252,978

EARLIER FILING DATE: 1994-06-02

NUMBER OF SEQ ID NOS: 34

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 28

LENGTH: 238

TYPE: PRT

ORGANISM: Staphylococcus aureas

US-08-896-933-28

Query Match  
Best Local Similarity 69.7%; Score 23; DB 4; Length 238;  
50.0%; Pred. No. 1.7;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
| | | | |  
DB 150 KKSVAQELD 159

RESULT 5  
US-09-314-235-28  
Sequence 28, Application US/09314235  
Patent No. 6338845  
GENERAL INFORMATION:

APPLICANT: Terman, David S.

TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,  
AND RELATED COMPOUNDS

FILE REFERENCE: 09629/005004

CURRENT APPLICATION NUMBER: US/09/314,235

EARLIER FILING DATE: 1999-05-18

EARLIER APPLICATION NUMBER: 08/896,933

EARLIER FILING DATE: 1997-07-18

EARLIER APPLICATION NUMBER: 08/252,978

EARLIER FILING DATE: 1994-06-02

EARLIER APPLICATION NUMBER: 07/891,718

EARLIER FILING DATE: 1992-06-01

EARLIER APPLICATION NUMBER: US91/00342

EARLIER FILING DATE: 1991-01-17

EARLIER APPLICATION NUMBER: 07/466,577

EARLIER FILING DATE: 1990-01-17

EARLIER APPLICATION NUMBER: 07/416,530

EARLIER FILING DATE: 1989-10-03

NUMBER OF SEQ ID NOS: 34

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 28

LENGTH: 238

TYPE: PRT

ORGANISM: Staphylococcus aureas

US-09-314-235-28

Query Match  
Best Local Similarity 69.7%; Score 23; DB 4; Length 238;  
50.0%; Pred. No. 1.7;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
| | | | |  
DB 150 KKSVAQELD 159

RESULT 6



US-08-896-933-27  
; Sequence 27, Application US/08896933  
; Patent No. 6221351  
; GENERAL INFORMATION:  
; APPLICANT: Terman, David S.  
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,  
; FILE OF INVENTION: SUPERANTIGENS, AND RELATED COMPOUNDS  
; FILE REFERENCE: 09629/005002  
; CURRENT APPLICATION NUMBER: US/08/896,933  
; CURRENT FILING DATE: 1997-07-18  
; EARLIER APPLICATION NUMBER: 08/252,978  
; EARLIER FILING DATE: 1994-06-02  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO: 27  
; LENGTH: 239  
; TYPE: PRF  
; ORGANISM: Staphylococcus aureas  
US-08-896-933-27

Query Match 69.7%; Score 23; DB 4; Length 239;  
Best Local Similarity 50.0%; Pred. No. 17;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEND 12  
| | | |  
Db 151 KKSVAQELD 160

RESULT 7  
US-09-314-235-27  
; Sequence 27, Application US/09314235.  
; Patent No. 6338845  
; GENERAL INFORMATION:  
; APPLICANT: Terman, David S.  
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,  
; FILE OF INVENTION: SUPERANTIGENS, AND RELATED COMPOUNDS  
; FILE REFERENCE: 09629/005004  
; CURRENT APPLICATION NUMBER: US/09/314,235  
; CURRENT FILING DATE: 1999-05-18  
; EARLIER APPLICATION NUMBER: 08/896,933  
; EARLIER FILING DATE: 1997-07-18  
; EARLIER APPLICATION NUMBER: 08/252,978  
; EARLIER FILING DATE: 1994-06-02  
; EARLIER APPLICATION NUMBER: 07/891,718  
; EARLIER FILING DATE: 1992-06-01  
; EARLIER APPLICATION NUMBER: US91/00342  
; EARLIER FILING DATE: 1991-01-17  
; EARLIER APPLICATION NUMBER: 07/466,577  
; EARLIER FILING DATE: 1990-01-17  
; EARLIER APPLICATION NUMBER: 07/416,530  
; EARLIER FILING DATE: 1989-10-03  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO: 27  
; LENGTH: 239  
; TYPE: PRF  
; ORGANISM: Staphylococcus aureas  
US-09-314-235-27

Query Match 69.7%; Score 23; DB 4; Length 239;  
Best Local Similarity 50.0%; Pred. No. 17;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEND 12  
| | | |  
Db 151 KKSVAQELD 160

RESULT 8  
US-09-144-776B-14  
; Sequence 14, Application US/09144776B  
; Patent No. 6399332

GENERAL INFORMATION:  
APPLICANT: Robert G. Ulrich,  
Mark A. Olson  
Sina Bavari  
TITLE OF INVENTION: Bacterial Superantigen  
Vaccines  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Charles H. Harris  
STREET: US Army MRC-504 Scott Street  
MCAR-JA (Charles H. Harris-Patent  
Atty)  
CITY: FORT DETRICK  
STATE: MARYLAND  
COUNTRY: USA  
ZIP: 21702-5012  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh 7.5  
SOFTWARE: Microsoft Word 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/144,776B  
FILING DATE: 01-Sep-1998  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/882,431  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Charles H. Harris  
REGISTRATION NUMBER: 34,616  
REFERENCE/DOCKET NUMBER: <Unknown>  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 619-2065  
TELEFAX: (301) 619-7714  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 266  
TYPE: Amino Acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 14:  
US-09-144-776B-14

Query Match 69.7%; Score 23; DB 4; Length 266;  
Best Local Similarity 50.0%; Pred. No. 19;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEND 12  
| | | |  
Db 178 KKSVAQELD 187

RESULT 9  
US-08-896-933-4  
; Sequence 4, Application US/08896933  
; Patent No. 6221351  
; GENERAL INFORMATION:  
; APPLICANT: Terman, David S.  
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,  
; FILE OF INVENTION: SUPERANTIGENS, AND RELATED COMPOUNDS  
; FILE REFERENCE: 09629/005002  
; CURRENT APPLICATION NUMBER: US/08/896,933  
; CURRENT FILING DATE: 1997-07-18  
; EARLIER APPLICATION NUMBER: 08/252,978  
; EARLIER FILING DATE: 1994-06-02  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO: 4  
; LENGTH: 17  
; TYPE: PRF  
; ORGANISM: Staphylococcus aureas

US-08-896-933-4

Query Match 66.7%; Score 22; DB 4; Length 17;  
Best Local Similarity 50.0%; Pred. No. 1.7;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12  
DB 1 KKNVTAQELD 10

RESULT 10

US-08-896-933-12  
Sequence 12, Application US/08896933  
Patent No. 6221351  
GENERAL INFORMATION:  
APPLICANT: Terman, David S.  
TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,  
FILE REFERENCE: 09629/005002  
CURRENT APPLICATION NUMBER: US/08/896,933  
CURRENT FILING DATE: 1997-07-18  
EARLIER APPLICATION NUMBER: 08/252,978  
EARLIER FILING DATE: 1994-06-02  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 12  
LENGTH: 17  
TYPE: PRT  
ORGANISM: Streptococcus pyogenes  
US-08-896-933-12

Query Match 66.7%; Score 22; DB 4; Length 17;  
Best Local Similarity 50.0%; Pred. No. 1.7;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12  
DB 1 KKNVTAQELD 10

RESULT 11

US-08-896-933-14  
Sequence 14, Application US/08896933  
Patent No. 6221351  
GENERAL INFORMATION:  
APPLICANT: Terman, David S.  
TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,  
FILE REFERENCE: 09629/005002  
CURRENT APPLICATION NUMBER: US/08/896,933  
CURRENT FILING DATE: 1997-07-18  
EARLIER APPLICATION NUMBER: 08/252,978  
EARLIER FILING DATE: 1994-06-02  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 14  
LENGTH: 17  
TYPE: PRT  
ORGANISM: Other  
FEATURE:  
OTHER INFORMATION: Consensus sequences derived from staphylococcus  
; OTHER INFORMATION: aureus and streptococcus pyogenes polypeptides  
US-08-896-933-14

Query Match 66.7%; Score 22; DB 4; Length 17;  
Best Local Similarity 50.0%; Pred. No. 1.7;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12  
DB 1 KKNVTAQELD 10

RESULT 12  
US-08-896-933-15  
Sequence 15, Application US/08896933  
Patent No. 6221351  
GENERAL INFORMATION:  
APPLICANT: Terman, David S.  
TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,  
FILE REFERENCE: 09629/005002  
CURRENT APPLICATION NUMBER: US/08/896,933  
CURRENT FILING DATE: 1997-07-18  
EARLIER APPLICATION NUMBER: 08/252,978  
EARLIER FILING DATE: 1994-06-02  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 15  
LENGTH: 17  
TYPE: PRT  
ORGANISM: Other  
FEATURE:  
OTHER INFORMATION: Consensus sequences derived from staphylococcus  
; OTHER INFORMATION: aureus and streptococcus pyogenes polypeptides  
US-08-896-933-15

Query Match 66.7%; Score 22; DB 4; Length 17;  
Best Local Similarity 50.0%; Pred. No. 1.7;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12  
DB 1 KKNVTAQELD 10

RESULT 13

US-08-896-933-16  
Sequence 16, Application US/08896933  
Patent No. 6221351  
GENERAL INFORMATION:  
APPLICANT: Terman, David S.  
TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,  
FILE REFERENCE: 09629/005002  
CURRENT APPLICATION NUMBER: US/08/896,933  
CURRENT FILING DATE: 1997-07-18  
EARLIER APPLICATION NUMBER: 08/252,978  
EARLIER FILING DATE: 1994-06-02  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 16  
LENGTH: 17  
TYPE: PRT  
ORGANISM: Other  
FEATURE:  
OTHER INFORMATION: Consensus sequences derived from staphylococcus  
; OTHER INFORMATION: aureus and streptococcus pyogenes polypeptides  
US-08-896-933-16

Query Match 66.7%; Score 22; DB 4; Length 17;  
Best Local Similarity 50.0%; Pred. No. 1.7;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12  
DB 1 KKNVTAQELD 10

RESULT 14

US-08-896-933-17  
Sequence 17, Application US/08896933  
Patent No. 6221351  
GENERAL INFORMATION:  
APPLICANT: Terman, David S.

;; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,  
;; TITLE OF INVENTION: SUPERANTIGENS, AND RELATED COMPOUNDS  
;; FILE REFERENCE: 09629/005002  
;; CURRENT APPLICATION NUMBER: US/08/896,933  
;; CURRENT FILING DATE: 1997-07-18  
;; EARLIER APPLICATION NUMBER: 08/252,978  
;; EARLIER FILING DATE: 1994-06-02  
;; NUMBER OF SEQ ID NOS: 34  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 17  
;; LENGTH: 17  
;; TYPE: PRT  
;; ORGANISM: Other  
;; FEATURE:  
;; OTHER INFORMATION: Consensus sequences derived from staphylococcus  
;; OTHER INFORMATION: aureus and streptococcus pyogenes polypeptides  
US-08-896-933-17

Query Match 66.7%; Score 22; DB 4; Length 17;  
Best Local Similarity 50.0%; Pred. No. 1.7;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTQXEND 12  
| | | |  
Db 1 KKNVTAQELD 10

RESULT 15  
US-09-314-235-4  
;; Sequence 4, Application US/09314235  
;; Patent No. 6338845  
;; GENERAL INFORMATION:  
;; APPLICANT: Terman, David S.  
;; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,  
;; TITLE OF INVENTION: SUPERANTIGENS, AND RELATED COMPOUNDS  
;; FILE REFERENCE: 09629/005004  
;; CURRENT APPLICATION NUMBER: US/09/314,235  
;; CURRENT FILING DATE: 1999-05-18  
;; EARLIER APPLICATION NUMBER: 08/896,933  
;; EARLIER FILING DATE: 1997-07-18  
;; EARLIER APPLICATION NUMBER: 08/252,978  
;; EARLIER FILING DATE: 1994-06-02  
;; EARLIER APPLICATION NUMBER: 07/891,718  
;; EARLIER FILING DATE: 1992-06-01  
;; EARLIER APPLICATION NUMBER: US91/00342  
;; EARLIER FILING DATE: 1991-01-17  
;; EARLIER APPLICATION NUMBER: 07/466,577  
;; EARLIER FILING DATE: 1990-01-17  
;; EARLIER APPLICATION NUMBER: 07/416,530  
;; EARLIER FILING DATE: 1989-10-03  
;; NUMBER OF SEQ ID NOS: 34  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 4  
;; LENGTH: 17  
;; TYPE: PRT  
;; ORGANISM: Staphylococcus aureas  
US-09-314-235-4

Query Match 66.7%; Score 22; DB 4; Length 17;  
Best Local Similarity 50.0%; Pred. No. 1.7;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTQXEND 12  
| | | |  
Db 1 KKNVTAQELD 10

Search completed: June 5, 2003, 15:32:44  
Job time : 11.5455 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:31:01 ; Search time 14.5455 Seconds  
(without alignments)  
85.173 Million cell updates/sec

Title: US-09-150-947f-16  
Perfect score: 33  
Sequence: 1 XXXXXXXXOEXD 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues  
Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA.\*  
1: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubppa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubppa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubppa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubppa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubppa/PCT05\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubppa/US08\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB.pep.\*  
9: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubppa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubppa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep.\*  
  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	69.7	29	10 US-09-864-761-43487	Sequence 43487, A
2	23	69.7	266	1 US-08-882-431-14	Sequence 14, Appl
3	23	69.7	266	9 US-09-870-759-12	Sequence 12, Appl
4	23	69.7	266	9 US-10-002-784A-14	Sequence 14, Appl
5	22	66.7	10	10 US-09-150-947B-3	Sequence 3, Appl1
6	22	66.7	10	10 US-09-150-947B-4	Sequence 4, Appl1
7	22	66.7	12	9 US-10-113-809-1	Sequence 1, Appl1
8	22	66.7	12	10 US-09-150-947B-2	Sequence 2, Appl1
9	22	66.7	13	10 US-09-150-947B-5	Sequence 5, Appl1
10	22	66.7	13	10 US-09-150-947B-6	Sequence 6, Appl1
11	22	66.7	13	10 US-09-150-947B-11	Sequence 11, Appl1
12	22	66.7	14	10 US-09-150-947B-9	Sequence 9, Appl1
13	22	66.7	14	10 US-09-150-947B-10	Sequence 10, Appl1
14	22	66.7	16	10 US-09-150-947B-7	Sequence 7, Appl1
15	22	66.7	26	10 US-09-150-947B-1	Sequence 1, Appl1
16	22	66.7	35	10 US-09-150-947B-8	Sequence 8, Appl1
17	22	66.7	102	9 US-10-047-260-26	Sequence 26, Appl
18	22	66.7	230	9 US-10-002-784A-26	Sequence 26, Appl

20	22	66.7	239	1 US-08-882-431-10	Sequence 10, Appl
21	22	66.7	239	9 US-10-002-784A-10	Sequence 10, Appl
22	22	66.7	239	10 US-09-150-947B-12	Sequence 12, Appl
23	22	66.7	251	1 US-08-882-431-16	Sequence 16, Appl
24	22	66.7	251	8 US-08-973-93A-13	Sequence 13, Appl
25	22	66.7	251	9 US-10-002-784A-16	Sequence 16, Appl
26	22	66.7	251	1 US-08-882-431-6	Sequence 6, Appl1
27	22	66.7	251	1 US-08-882-431-8	Sequence 8, Appl1
28	22	66.7	266	9 US-09-870-759-10	Sequence 10, Appl
29	22	66.7	266	9 US-10-002-784A-6	Sequence 6, Appl1
30	22	66.7	266	9 US-10-002-784A-8	Sequence 8, Appl1
31	22	66.7	266	9 US-10-151-336-8	Sequence 8, Appl1
32	22	66.7	295	10 US-09-815-242-13351	Sequence 13351, A
33	22	66.7	295	10 US-09-815-242-13691	Sequence 13691, A
34	22	66.7	344	9 US-10-201-310-3	Sequence 3, Appl1
35	22	66.7	350	9 US-09-972-473-38	Sequence 38, Appl
36	22	66.7	468	9 US-10-002-784A-27	Sequence 27, Appl
37	22	66.7	659	9 US-10-102-627-74	Sequence 74, Appl
38	22	66.7	922	10 US-09-879-228-2	Sequence 2, Appl1
39	22	63.6	12	9 US-10-113-809-2	Sequence 2, Appl1
40	21	63.6	72	9 US-10-144-929-189	Sequence 189, App
41	21	63.6	177	9 US-10-092-849-5	Sequence 5, Appl1
42	21	63.6	177	10 US-09-730-617-95	Sequence 95, Appl
43	21	63.6	180	9 US-09-738-226-4637	Sequence 4637, Ap
44	21	63.6	203	9 US-09-900-766-5	Sequence 5, Appl1
45	21	63.6	217	9 US-09-900-766-6	Sequence 6, Appl1

## ALIGNMENTS

RESULT 1  
US-09-864-761-43487  
Sequence 43487, Application US/09864761  
Patient No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wenshang  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO  
FILE REFERENCE: Neomulca-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 43487  
LENGTH: 29  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC006299.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.8  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.7  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.67  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.7  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.66  
OTHER INFORMATION: EST\_HUMAN HIT: AV703706.1, EVALU6 6.00e+00  
US-09-864-761-43487

Query Match 69.7%; Score 23; DB 10; Length 29;  
Best Local Similarity 50.0%; Pred. No. 6.8;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
| | | |  
DB 10 KALSTQGEID 19

RESULT 2  
US-08-882-431-14  
Sequence 14, Application US/08882431  
Publication No. US20030009015A1  
GENERAL INFORMATION:  
APPLICANT: Robert G. Ulrich,  
APPLICANT: Mark A. Olson  
APPLICANT: Sina Bavari  
TITLE OF INVENTION: Bacterial Superantigen  
TITLE OF INVENTION: Vaccines  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John Moran  
STREET: US Army MRC -504 Scott Street MCMR-JA (John Moran-Patent Atty)  
CITY: FORT DETRICK  
STATE: MARYLAND  
COUNTRY: USA  
ZIP: 21702-5012  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh 7.5  
SOFTWARE: Microsoft Word 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/882,431  
FILING DATE: June 25, 1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Moran, John  
REGISTRATION NUMBER: 26,313  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 619-2065  
TELEFAX: (301) 619-7714  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 266

TYPE: Amino Acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
US-08-882-431-14

Query Match 69.7%; Score 23; DB 1; Length 266;  
Best Local Similarity 50.0%; Pred. No. 71;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
| | | |  
DB 178 KKSVTQGEID 187

RESULT 3  
US-09-870-759-12  
Sequence 12, Application US/09870759  
Patent No. US2002017751A1  
GENERAL INFORMATION:  
APPLICANT: TERMAN, David S  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE  
FILE REFERENCE: 870759  
CURRENT APPLICATION NUMBER: US/09/870,759  
CURRENT FILING DATE: 2002-01-14  
PRIOR APPLICATION NUMBER: US 60/208,128  
PRIOR FILING DATE: 2000-05-30  
NUMBER OF SEQ ID NOS: 166  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 12  
LENGTH: 266  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-870-759-12

Query Match 69.7%; Score 23; DB 9; Length 266;  
Best Local Similarity 50.0%; Pred. No. 71;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
| | | |  
DB 178 KKSVTQGEID 187

RESULT 4  
US-10-002-784A-14  
Sequence 14, Application US/10002784A  
Publication No. US20030036644A1  
GENERAL INFORMATION:  
/33  
APPLICANT: Ulrich, Robert G.  
TITLE OF INVENTION: Bacterial Superantigen Vaccines  
FILE REFERENCE: 003/233/SAP  
CURRENT APPLICATION NUMBER: US/10/002,784A  
CURRENT FILING DATE: 2001-11-26  
PRIOR APPLICATION NUMBER: 08/882,431; 09/144,776  
PRIOR FILING DATE: 97-06-25; 98-09-01  
NUMBER OF SEQ ID NOS: 40  
SOFTWARE: Apple Macintosh Microsoft Word 6.0  
SEQ ID NO 14  
LENGTH: 266  
TYPE: PRT  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: staphylococcal enterotoxin C-1 mutant  
US-10-002-784A-14

Query Match 69.7%; Score 23; DB 9; Length 266;  
Best Local Similarity 50.0%; Pred. No. 71;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
| | | |

Db 178 KKSVTAQELD 187

## RESULT 5

US-09-150-947B-3

; Sequence 3, Application US/09150947B  
; Patent No. US20020028211A1

; GENERAL INFORMATION:

; APPLICANT: Kaempfer, Raymond

; APPLICANT: Arad, Gila

; TITLE OF INVENTION: BROAD SPECTRUM PYROGENIC EXOTOXINS

; FILE REFERENCE: A31967-PCT-USA-A

; CURRENT APPLICATION NUMBER: US/09/150,947B

; CURRENT FILING DATE: 1998-09-10

; PRIOR APPLICATION NUMBER: PCT/IL97/00438

; PRIOR FILING DATE: 1997-12-30

; PRIOR APPLICATION NUMBER: ISRAEL 119938

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 3

; LENGTH: 10

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-150-947B-3

US-09-150-947B-3

Query Match 66.7%; Score 22; DB 10; Length 10;

Best Local Similarity 50.0%; Pred. No. 4.2;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12

Db 1 KKKVTAQELD 10

## RESULT 6

US-09-150-947B-4

; Sequence 4, Application US/09150947B  
; Patent No. US20020028211A1

; GENERAL INFORMATION:

; APPLICANT: Kaempfer, Raymond

; APPLICANT: Arad, Gila

; TITLE OF INVENTION: BROAD SPECTRUM PYROGENIC EXOTOXINS

; FILE REFERENCE: A31967-PCT-USA-A

; CURRENT APPLICATION NUMBER: US/09/150,947B

; CURRENT FILING DATE: 1998-09-10

; PRIOR APPLICATION NUMBER: PCT/IL97/00438

; PRIOR FILING DATE: 1997-12-30

; PRIOR APPLICATION NUMBER: ISRAEL 119938

; PRIOR FILING DATE: 1996-12-30

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 4

; LENGTH: 10

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-150-947B-4

Query Match 66.7%; Score 22; DB 10; Length 10;

Best Local Similarity 50.0%; Pred. No. 4.2;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12

Db 1 KKKVTAQELD 10

## RESULT 7

US-10-113-809-1

; Sequence 1, Application US/10113809  
; Patent No. US20020177573A1

; GENERAL INFORMATION:

; APPLICANT: Masche, Hans, E.

; TITLE OF INVENTION: EXOTOXIN-LIGAND

; FILE REFERENCE: MBP-011XX

; CURRENT APPLICATION NUMBER: US/10/113,809

; CURRENT FILING DATE: 2002-03-29

; PRIOR APPLICATION NUMBER: 101 16 042.9-41

; PRIOR FILING DATE: 2001-03-30

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 12

; TYPE: PRT

; ORGANISM: Artificial Sequence

; OTHER INFORMATION: Bacterial Toxin Binding Ligand

US-10-113-809-1

Query Match 66.7%; Score 22; DB 9; Length 12;

Best Local Similarity 50.0%; Pred. No. 5.1;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12

Db 3 KKKVTAQELD 12

## RESULT 8

US-10-113-809-4

; Sequence 4, Application US/10113809  
; Patent No. US20020177573A1

; GENERAL INFORMATION:

; APPLICANT: Masche, Hans, E.

; TITLE OF INVENTION: EXOTOXIN-LIGAND

; FILE REFERENCE: MBP-011XX

; CURRENT APPLICATION NUMBER: US/10/113,809

; CURRENT FILING DATE: 2002-03-29

; PRIOR APPLICATION NUMBER: 101 16 042.9-41

; PRIOR FILING DATE: 2001-03-30

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 12

; TYPE: PRT

; ORGANISM: Artificial Sequence

; OTHER INFORMATION: Bacterial Toxin Binding Ligand

US-10-113-809-4

Query Match 66.7%; Score 22; DB 9; Length 12;

Best Local Similarity 50.0%; Pred. No. 5.1;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12

Db 3 KKKVTAQELD 12

## RESULT 9

US-09-150-947B-2

; Sequence 2, Application US/09150947B  
; Patent No. US20020028211A1

; GENERAL INFORMATION:

; APPLICANT: Kaempfer, Raymond

; APPLICANT: Arad, Gila

; TITLE OF INVENTION: BROAD SPECTRUM PYROGENIC EXOTOXINS

; FILE REFERENCE: A31967-PCT-USA-A

; CURRENT APPLICATION NUMBER: US/09/150,947B

; CURRENT FILING DATE: 1998-09-10

; PRIOR APPLICATION NUMBER: PCT/IL97/00438

; PRIOR FILING DATE: 1997-12-30

; PRIOR APPLICATION NUMBER: ISRAEL 119938

```
; PRIOR FILING DATE: 1996-12-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-150-947B-2
```

```
Query Match      66.7%; Score 22; DB 10; Length 12;
Best Local Similarity 50.0%; Pred. No. 5.1;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      3 KXXXTXQEXD 12
        | | | |
DB      4 KKRATVQELD 13
```

```
RESULT 10
US-09-150-947B-5
; Sequence 5, Application US/09150947B
; Patent No. US20020028211A1
; GENERAL INFORMATION:
; APPLICANT: Kaempfer, Raymond
; APPLICANT: Arad, Gila
; TITLE OF INVENTION: BROAD SPECTRUM PYROGENIC EXOTOXINS
; FILE REFERENCE: A31967-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/150,947B
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: PCT/IL97/00438
; PRIOR FILING DATE: 1997-12-30
; PRIOR APPLICATION NUMBER: ISRAEL 119938
; PRIOR FILING DATE: 1996-12-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: LIPID
; LOCATION: (1)...(1)
; OTHER INFORMATION: N-lauryl cysteine residue
US-09-150-947B-5
```

```
Query Match      66.7%; Score 22; DB 10; Length 13;
Best Local Similarity 50.0%; Pred. No. 5.5;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      3 KXXXTXQEXD 12
        | | | |
DB      4 KKRATVQELD 13
```

```
RESULT 11
US-09-150-947B-6
; Sequence 6, Application US/09150947B
; Patent No. US20020028211A1
; GENERAL INFORMATION:
; APPLICANT: Kaempfer, Raymond
; APPLICANT: Arad, Gila
; TITLE OF INVENTION: BROAD SPECTRUM PYROGENIC EXOTOXINS
; FILE REFERENCE: A31967-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/150,947B
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: PCT/IL97/00438
; PRIOR FILING DATE: 1997-12-30
; PRIOR APPLICATION NUMBER: ISRAEL 119938
; PRIOR FILING DATE: 1996-12-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 6
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: LIPID
; LOCATION: (1)...(1)
; OTHER INFORMATION: N-lauryl cysteine residue
US-09-150-947B-6
```

```
Query Match      66.7%; Score 22; DB 10; Length 13;
Best Local Similarity 50.0%; Pred. No. 5.5;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      3 KXXXTXQEXD 12
        | | | |
DB      4 KKRATVQELD 13
```

```
RESULT 12
US-09-150-947B-11
; Sequence 11, Application US/09150947B
; Patent No. US20020028211A1
; GENERAL INFORMATION:
; APPLICANT: Kaempfer, Raymond
; APPLICANT: Arad, Gila
; TITLE OF INVENTION: BROAD SPECTRUM PYROGENIC EXOTOXINS
; FILE REFERENCE: A31967-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/150,947B
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: PCT/IL97/00438
; PRIOR FILING DATE: 1997-12-30
; PRIOR APPLICATION NUMBER: ISRAEL 119938
; PRIOR FILING DATE: 1996-12-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: ACETYLATION
; LOCATION: (1)...(1)
; OTHER INFORMATION: N-acetyl group
; NAME/KEY: MOD.RES
; LOCATION: (13)...(13)
; OTHER INFORMATION: D-alanine
US-09-150-947B-11
```

```
Query Match      66.7%; Score 22; DB 10; Length 13;
Best Local Similarity 50.0%; Pred. No. 5.5;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      3 KXXXTXQEXD 12
        | | | |
DB      4 KKRATVQELD 13
```

```
RESULT 13
US-09-150-947B-9
; Sequence 9, Application US/09150947B
; Patent No. US20020028211A1
; GENERAL INFORMATION:
; APPLICANT: Kaempfer, Raymond
; APPLICANT: Arad, Gila
; TITLE OF INVENTION: BROAD SPECTRUM PYROGENIC EXOTOXINS
; FILE REFERENCE: A31967-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/150,947B
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: PCT/IL97/00438
; PRIOR FILING DATE: 1997-12-30
```



```

; PRIOR APPLICATION NUMBER: ISRAEL 119938
; PRIOR FILING DATE: 1996-12-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: DISULFID
; LOCATION: (1)...(1)
; NAME/KEY: DISULFID
; LOCATION: (14)...(14)
US-09-150-947B-9
```

```

Query Match      66.7%; Score 22; DB 10; Length 14;
Best Local Similarity 50.0%; Pred. No. 6;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      3 KXXXTXQEXD 12
      | | | | |
DB      4 KKKATVQELD 13
```

```

RESULT 14
US-09-150-947B-10
; Sequence 10, Application US/09150947B
; Patent No. US2002002821A1
; GENERAL INFORMATION:
; APPLICANT: Kaempfer, Raymond
; APPLICANT: Arad, Gila
; TITLE OF INVENTION: BROAD SPECTRUM PYROGENIC EXOTOXINS
; FILE REFERENCE: A31967-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/150,947B
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: PCT/IL97/00438
; PRIOR FILING DATE: 1997-12-30
; PRIOR APPLICATION NUMBER: ISRAEL 119938
; PRIOR FILING DATE: 1996-12-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (1)...(1)
; OTHER INFORMATION: D-alanine
; NAME/KEY: MOD_RES
; LOCATION: (14)...(14)
; OTHER INFORMATION: D-alanine
US-09-150-947B-10
```

```

Query Match      66.7%; Score 22; DB 10; Length 14;
Best Local Similarity 50.0%; Pred. No. 6;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      3 KXXXTXQEXD 12
      | | | | |
DB      4 KKKATVQELD 13
```

```

RESULT 15
US-09-150-947B-7
; Sequence 7, Application US/09150947B
; Patent No. US2002002821A1
; GENERAL INFORMATION:
; APPLICANT: Kaempfer, Raymond
; APPLICANT: Arad, Gila
; TITLE OF INVENTION: BROAD SPECTRUM PYROGENIC EXOTOXINS
; TITLE OF INVENTION: ANTAGONISTS AND VACCINES
```

```

; FILE REFERENCE: A31967-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/150,947B
; CURRENT FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: PCT/IL97/00438
; PRIOR FILING DATE: 1997-12-30
; PRIOR APPLICATION NUMBER: ISRAEL 119938
; PRIOR FILING DATE: 1996-12-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-150-947B-7
```

```

Query Match      66.7%; Score 22; DB 10; Length 16;
Best Local Similarity 50.0%; Pred. No. 6.9;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      3 KXXXTXQEXD 12
      | | | | |
DB      3 KKKATVQELD 12
```

Search completed: June 5, 2003, 15:49:05  
Job time : 14.5455 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:24:01 ; Search time 11.8182 Seconds  
(without alignments)  
97.613 Million cell updates/sec

Title: US-09-150-947f-16  
Perfect score: 33  
Sequence: 1 XXXXXXTXQEXD 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	72.7	549	2 E86337	hypothetical prote
2	24	72.7	602	1 RRMZP3	polymerase-associ
3	24	72.7	603	1 RRMZP5	polymerase-associ
4	24	72.7	604	1 RRMZP4	polymerase-associ
5	23	69.7	266	1 ENSAC1	enterotoxin C-1 pr
6	23	69.7	266	2 A60114	enterotoxin C-2 - S
7	23	69.7	266	2 S11885	enterotoxin C-3 - S
8	23	69.7	565	2 A87694	ribosomal protein
9	23	69.7	566	2 AB3061	30S ribosomal prot
10	23	69.7	568	1 R3ZRL1	ribosomal protein
11	23	69.7	572	2 E98225	30S ribosomal prot
12	23	69.7	586	1 ORXLE	estrogen receptor
13	23	69.7	3212	2 T24692	hypothetical prote
14	22	66.7	102	2 S74843	hypothetical prote
15	22	66.7	136	2 A89969	enterotoxin YEMT2
16	22	66.7	184	2 F86681	prophage pII prote
17	22	66.7	236	2 S18789	exotoxin A precurs
18	22	66.7	236	2 S18786	exotoxin type A pr
19	22	66.7	236	2 S18783	exotoxin type A pr
20	22	66.7	239	2 D89969	enterotoxin SEM 11
21	22	66.7	240	2 G89991	extracellular ente
22	22	66.7	242	2 C89969	extracellular ente
23	22	66.7	247	2 T22535	hypothetical prote
24	22	66.7	251	2 S29659	exotoxin type A pr
25	22	66.7	251	1 QOECF	replication initia
26	22	66.7	258	2 H89968	enterotoxin Sen 11
27	22	66.7	260	2 E89969	enterotoxin Sen 11
28	22	66.7	266	1 ENSAB6	enterotoxin B prec
29	22	66.7	295	2 D95107	hypothetical prote

30	22	66.7	295	2 F97975	hypothetical prote
31	22	66.7	332	1 DEBYG2	glyceraldhyde-3-P
32	22	66.7	333	1 JCS787	ribonuclease H (EC
33	22	66.7	356	2 T38408	hypothetical prote
34	22	66.7	520	2 G71647	hypothetical prote
35	22	66.7	523	2 F97865	hypothetical prote
36	22	66.7	672	2 F71424	hypothetical prote
37	22	66.7	835	2 T05259	hypothetical prote
38	22	66.7	886	2 T39229	probable disease r
39	22	66.7	922	2 T37292	hypothetical prote
40	22	66.7	946	2 A84133	DNA ligase (ATP) (
41	22	66.7	1146	2 A86796	hypothetical prote
42	22	66.7	1200	2 S77524	chromosome segrega
43	22	66.7	1314	2 A85176	hypothetical prote
44	22	66.7	1327	2 D70759	probable ctsb prot
45	22	66.7	1738	2 T14867	interaptin - slime

## ALIGNMENTS

## RESULT 1

E86337 hypothetical protein F14010.3 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001

C:Accession: E86337

R:Rheologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon  
Chin, C.W.; Chung, M.K.; Comp, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,  
ansen, N.F.; Hughes, B.; Hultzer, L.  
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luroo, J.S.; Maitl, R.; Marzia  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shun, P.; Southwick, A.M.; Sun, H.; Tallo  
ker, M.; Wu, D.; Xu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: E86337

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-549 <STO>

A:Cross-references: GB:AE005172; NID:95558589; PIDN:AAF8152.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

Query Match 72.7% Score 24; DB 2; Length 549;  
Best Local Similarity 50.0% Pred. No. 62;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 3 XXXXXQEXD 12  
Db 27 KISTGTQEMD 36

## RESULT 2

RRMZP3 polymerase-associated nucleocapsid phosphoprotein - parainfluenza virus type 3 (strat

C:Species: parainfluenza virus type 3

A:Note: host Homo sapiens (man)  
C>Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 24-Jul-1997

C:Accession: A27010

R:SPR19gs, M.K.; Collins, P.L.

J. Gen. Virol. 67, 2705-2719, 1986

A:Title: Sequence analysis of the P and C protein genes of human parainfluenza virus

A:Reference number: A92793; MUID:87085488; PMID:3025346

A:Accession: A27010

A:Molecule type: mRNA

A:Residues: 1-602 <SPR>

C:Comment: This protein may be a component of the active polymerase.

C:Genetics:

A:Gene: P

C:Superfamily: parainfluenza virus polymerase-associated nucleocapsid phosphoprotein

C:Keywords: nucleocapsid; phosphoprotein

Query Match 72.7%; Score 24; DB 1; Length 602;  
Best Local Similarity 50.0%; Pred. No. 68;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
| | | | |  
DB 226 KSSSTHQEDD 235

#### RESULT 3

RRNZP5

polymerase-associated nucleocapsid phosphoprotein (version 2) - parainfluenza virus type C:Species: parainfluenza virus type 3

A:Note: host Homo sapiens (man)

C:Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 24-Jul-1997

C:Accession: A26896

R:Gallinetti, M.S.; Mink, M.A.; Lambert, D.M.; Wechsler, S.L.; Pons, M.W.

Virology 155, 46-60, 1986

A:Title: Molecular cloning and sequence analysis of the human parainfluenza 3 virus mRNA

A:Reference number: A94348; PMID:87044104; PMID:3022477

A:Accession: A26896

A:Molecule type: mRNA

A:Residues: 1-603 <GAL>

C:Comment: The RNA sequence was obtained from GenBank, release 52.0.

C:Genetics:

A:Gene: P

C:Superfamily: parainfluenza virus polymerase-associated nucleocapsid phosphoprotein

C:Keywords: nucleocapsid; phosphoprotein

Query Match 72.7%; Score 24; DB 1; Length 603;  
Best Local Similarity 50.0%; Pred. No. 68;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
| | | | |  
DB 226 KSSSTHQEDD 235

#### RESULT 4

RRNZP4

polymerase-associated nucleocapsid phosphoprotein (version 1) - parainfluenza virus type C:Species: parainfluenza virus type 3

A:Note: host Homo sapiens (man)

C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 24-Jul-1997

C:Accession: A94355; A24189

R:Luik, D.; Sanchez, A.; Banerjee, A.K.

Virology 156, 193-194, 1987

A:Reference number: A94355

A:Contents: extratum

A:Accession: A94355

A:Molecule type: mRNA

A:Residues: 1-604 <LUI>

A:Cross-references: GB:M14890

R:Luik, D.; Sanchez, A.; Banerjee, A.K.

Virology 153, 318-325, 1986

A:Title: Messenger RNA encoding the phosphoprotein (P) gene of human parainfluenza virus

A:Reference number: A94343; PMID:86291173; PMID:3016995

A:Accession: A24189

A:Molecule type: mRNA

A:Residues: 1-603, 'AKDQIKKTPNK' <LUI>

A:Note: this sequence has been corrected in reference A94355

C:Comment: This protein may be a component of the active polymerase.

C:Genetics:

A:Gene: P

C:Superfamily: parainfluenza virus polymerase-associated nucleocapsid phosphoprotein

C:Keywords: nucleocapsid; phosphoprotein

Query Match 72.7%; Score 24; DB 1; Length 604;  
Best Local Similarity 50.0%; Pred. No. 69;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
| | | | |  
DB 226 KSSSTHQEDD 235

#### RESULT 5

ENSAC1

enterotoxin C-1 precursor - Staphylococcus aureus C:Species: Staphylococcus aureus

C:Date: 15-Nov-1984 #sequence\_revision 05-Jan-1996 #text\_change 18-Jun-1999

C:Accession: S06356; A01816

R:Bohach, G.A.; Schlievert, P.M.

Mol. Gen. Genet. 209, 15-20, 1987

A:Title: Nucleotide sequence of the staphylococcal enterotoxin C1 gene and relatednes

A:Reference number: S06356; PMID:88038352; PMID:2823067

A:Accession: S06356

A:Molecule type: DNA

A:Residues: 1-266 <BOH>

A:Cross-references: EMBL:X05815; NID:g46566; PIDN:CAA29260.1; PID:g46567

R:Schmidt, J.J.; Spero, L.

J. Biol. Chem. 258, 6300-6306, 1983

A:Title: The complete amino acid sequence of Staphylococcal enterotoxin C-1.

A:Reference number: A01816; PMID:83213327; PMID:6189824

A:Molecule type: protein

A:Residues: 28-75, 'IL', '78-176, 'N', '178-266 <SCH>

C:Genetics:

A:Gene: entC1

C:Superfamily: enterotoxin B

C:Keywords: enterotoxin

F:1-27/Domain: signal sequence #status predicted <SIG>

F:28-266/Product: enterotoxin C-1 #status experimental <MAT>

F:120-137/Disulfide bonds: #status experimental

Query Match 69.7%; Score 23; DB 1; Length 266;  
Best Local Similarity 50.0%; Pred. No. 55;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
| | | | |  
DB 178 KSVTAQEDD 187

#### RESULT 6

A60114

enterotoxin C-2 precursor - Staphylococcus aureus N:Alternate names: enterotoxin C-3 precursor

C:Species: Staphylococcus aureus

C:Date: 10-Nov-1992 #sequence\_revision 10-Nov-1992 #text\_change 16-Jul-1999

C:Accession: A60114; B60114; A33866

R:Bohach, G.A.; Schlievert, P.M.

Infect. Immun. 57, 2249-2252, 1989

A:Title: Conservation of the biologically active portions of staphylococcal enterotox

A:Reference number: A60114; PMID:8927549; PMID:2543637

A:Accession: A60114

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-266 <BOH>

A:Accession: B60114

A:Molecule type: protein

A:Residues: 28-66 <BOH2>

R:Couch, J.L.; Betley, M.J.

J. Bacteriol. 171, 4507-4510, 1989

A:Title: Nucleotide sequence of the type C-3 staphylococcal enterotoxin gene suggests

A:Reference number: A33866; PMID:89327174; PMID:2473979

A:Accession: A33866

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-266 <COU>

A:Cross-references: GB:M28364; NID:g153003; PIDN:AAA26624.1; PID:g153004

C:Genetics:

A:Gene: entC2

C;Superfamily: enterotoxin B  
F:1-27/Domain: signal sequence #status predicted <SIG>  
F:28-266/Product: enterotoxin C-2 #status predicted <MAT>

Query Match 69.7%; Score 23; DB 2; Length 266;  
Best Local Similarity 50.0%; Pred. No. 55;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
DB 178 KKSXTAQEED 187

## RESULT 7

S11885  
enterotoxin C3 - Staphylococcus aureus

C;Species: Staphylococcus aureus  
C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 16-Jul-1999

C;Accession: S11885  
R;Hovde, C.J.; Hackelt, S.P.; Bohach, G.A.

Mol. Gen. Genet. 220, 329-333, 1990

A;Title: Nucleotide sequence of the staphylococcal enterotoxin C3 gene: sequence comparison  
A;Reference number: S11885; MID:90220508; PMID:2325627

A;Accession: S11885

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-266 <NOV>

A;Cross-references: GB:X51661; NID:946570; PIDN:CAA35972.1; PID:946571

C;Superfamily: enterotoxin B

Query Match 69.7%; Score 23; DB 2; Length 266;  
Best Local Similarity 50.0%; Pred. No. 55;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
DB 178 KKSXTAQEED 187

## RESULT 8

A87694  
ribosomal protein S1 [Imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus  
C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001

C;Accession: A87694  
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; Debey, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of Caulobacter crescentus.

A;Reference number: A87249; MID:21173698; PMID:11259647

A;Accession: A87694

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-565 <STO>

A;Cross-references: GB:AE005673; NID:913425331; PIDN:AAK2549.1; GSPDB:GN00148

C;Genetics:

A;Gene: CC3587

C;Superfamily: Escherichia coli ribosomal protein S1

Query Match 69.7%; Score 23; DB 2; Length 565;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
DB 326 KIVSTQSEVD 335

## RESULT 9

AB3061  
30S ribosomal protein S1 [Imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
C;Species: Agrobacterium tumefaciens

C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 01-Feb-2002  
C;Accession: AB3061  
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCl  
Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam  
ster, E.W.

A;Reference number: AB2577; PMID:11743193

A;Accession: AB3061

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-566 <KUR>

A;Cross-references: GB:AE008689; PIDN:AA44904.1; PID:917742555; GSPDB:GN00187

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: rpsA

A;Map position: linear chromosome

C;Superfamily: Escherichia coli ribosomal protein S1

Query Match 69.7%; Score 23; DB 2; Length 566;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
DB 326 KIVSTQSEVD 335

## RESULT 10

R32R1  
ribosomal protein S1 - Rhizobium meliloti

C;Species: Rhizobium meliloti  
C;Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 22-Jun-1999

C;Accession: S01055

R;Schneider, J.; Thamm, S.; Lutz, R.; Hussain, A.; Palst, G.; Dobrinski, B.

Nucleic Acids Res. 16, 3075-3089, 1988

A;Title: Cloning and characterization of a gene from Rhizobium meliloti 2011 coding

A;Reference number: S01055; MID:88217521; PMID:3368316

A;Accession: S01055

A;Molecule type: DNA

A;Residues: 1-568 <SCCH>

A;Cross-references: EMBL:X07528; NID:946338; PIDN:CAA30404.1; PID:946339

A;Note: the sequence from Fig. 5 is inconsistent with that from Fig. 4 in having 565-

C;Genetics:

A;Gene: rpsA

C;Superfamily: Escherichia coli ribosomal protein S1

C;Keywords: duplication; protein biosynthesis; ribosome; RNA binding

F:196-453/Domain: RNA binding #status predicted <RNA>

F:196-279,280-366,367-453/Region: duplication

Query Match 69.7%; Score 23; DB 1; Length 568;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
DB 326 KIVSTQSEVD 335

## RESULT 11

E98225  
30S ribosomal protein S1 [Imported] - Agrobacterium tumefaciens (strain C58, Cereon)

C;Species: Agrobacterium tumefaciens  
C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 11-Jan-2002

C;Accession: E98225

R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,

Science 294, 2323-2328, 2001

A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium

A;Reference number: A97359; PMID:11743194

A;Accession: E98225

A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-572 <KUR>  
A:Cross-references: GB:AE007670; PIDN:AAK89327.1; PID:g15159167; GSPDB:GN00170  
C:Genetics:  
A:Gene: AGR\_L1496  
A:Map position: linear chromosome  
C:Superfamily: Escherichia coli ribosomal protein S1

Query Match 69.7%; Score 23; DB 2; Length 572;  
Best Local Similarity 50.0%; Pred. No. 12e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12  
DB 332 KILSTQGEVD 341

## RESULT 12

ORXLE

estrogen receptor - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C>Date: 28-Feb-1992 #sequence\_revision 14-Jul-1994 #text\_change 15-Jun-1996  
C:Accession: M40907

R:Weller, I.J.; Lew, D.; Shapiro, D.J.  
MOL. Endocrinol. 1, 355-362, 1987

A:Title: The Xenopus laevis estrogen receptor: sequence homology with human and avian re  
A:Reference number: M40907; MUID:90331927; PMID:3274894  
A:Accession: M40907

A:Molecule type: mRNA

A:Residues: 1-586 <WEI>

A:Cross-references: GB:L20735

C:Comment: The steroid hormones and their receptors are involved in the regulation of eu

C:Comment: In the absence of ligand, steroid hormone receptors are thought to be weakly

C:Superfamily: estrogen receptor; erba transforming protein homology

C:Keywords: DNA binding; nucleus; phosphoprotein; steroid binding; steroid hormone recep

F:1-174/Domain: amino-terminal <NH2>

F:175-266/Domain: DNA binding #status predicted <DNA>

F:178-448/Domain: erba transforming protein homology <ERBA>

F:178-201/Region: zinc finger CCCC motif

F:214-237/Region: zinc finger CCCC motif

F:251-266/Region: nuclear location signal

F:292-544/Domain: steroid binding #status predicted <STR>

F:180,183,197,200/Binding site: zinc (Cys) #status predicted

F:215,222,232,235/Binding site: zinc (Cys) #status predicted

F:231/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match

Best Local Similarity 59.7%; Score 23; DB 1; Length 586;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12  
DB 545 KDKTTQGEVD 554

## RESULT 13

T24692

hypothetical protein T08G11.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T24692

R:Dobson, R.  
submitted to the EMBL Data Library, September 1996

A:Reference number: 219925

A:Accession: T24692

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-3212 <WIL>

A:Cross-references: EMBL:Z80220; PIDN:CA802304.1; GSPDB:GN00019; CESP:T08G11.1

A:Experimental source: clone T08G11

C:Genetics:

A:Gene: CESP:T08G11.1  
A:Map position: 1  
A:Introns: 34/1; 149/3; 182/3; 527/2; 766/1; 1070/3; 1769/3; 1834/3; 2070/2; 2210/3;

Query Match 69.7%; Score 23; DB 2; Length 3212;  
Best Local Similarity 50.0%; Pred. No. 8e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12  
DB 2264 KSVTTSQEND 2273

## RESULT 14

S74843

hypothetical protein s110846 - Synechocystis sp. (strain PCC 6803)

C:Species: Synechocystis sp.

A:Variety: PCC 6803

C>Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999

C:Accession: S74843

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys

S.

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S74843

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-102 <KAN>

A:Cross-references: EMBL:D90909; GB:AB001339; NID:g1652844; PIDN:BA17804.1; PID:g101

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 66.7%; Score 22; DB 2; Length 102;  
Best Local Similarity 50.0%; Pred. No. 38;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12  
DB 72 KKKTKQGEVD 81

## RESULT 15

A89969

enterotoxin YENT2 [Imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001

C:Accession: A89969

R:Kuruda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O

ma, A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsugu, K.

Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: A89969

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-136 <KUR>

A:Cross-references: GB:BA000018; PID:g13701619; PIDN:BA842912.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: yent2

Query Match 66.7%; Score 22; DB 2; Length 136;  
Best Local Similarity 50.0%; Pred. No. 51;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12  
DB 50 KKNITQGEVD 59

Search completed: June 5, 2003, 15:34:00

Job time : 11.8162 secs



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:34:07 ; Search time 7.81818 Seconds

(without alignments)  
63.661 Million cell updates/sec

Title: US-09-150-947F-16

Perfect score: 33

Sequence: 1 XXXXXXXXQEXD 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	72.7	604	RRP_P13H4	P06162 human para
2	69.7	266	ETC1_STAU	P01353 staphylococ
3	69.7	266	ETC2_STAU	P34071 staphylococ
4	69.7	266	ETC3_STAU	P23313 staphylococ
5	69.7	568	RSL_RHIME	P14129 rhizobium m
6	69.7	586	ESR1_XENLA	P81559 xenopus lae
7	66.7	251	REEL_ECOLI	P03856 escherichia
8	66.7	251	SPER_STRPY	P08095 streptococ
9	66.7	266	ETXB_STAU	P01552 staphylococ
10	66.7	331	GP3_YEAS	P00359 saccharomyc
11	66.7	350	DKK3_CHICK	P09839 gallus gall
12	66.7	356	YDGC_SCHPO	P10498 schizosacch
13	66.7	520	Y855_RICPR	P09245 rickettsia
14	66.7	712	ARS2_CAEEL	P09613 caenorhabd
15	66.7	886	YEB6_SCHPO	P014302 schizosacch
16	66.7	922	DNL3_HUMAN	P491916 homo sapien
17	66.7	1015	YK03_MOUSE	P97386 mus musculu
18	66.7	1327	YR59_LISMO	P08583 listeria mo
19	63.6	176	GL1_HORSE	P127900 arabidopsi
20	63.6	177	GL1_ARATH	P27900 arabidopsi
21	63.6	228	GL1_ARATH	P27900 arabidopsi
22	63.6	234	SPG_STRPY	P13380 streptococ
23	63.6	235	SPG_STRPY	P13380 streptococ
24	63.6	236	SPH_STRPY	P13380 streptococ
25	63.6	257	ETXA_STAU	P13363 staphylococ
26	63.6	257	ETXA_STAU	P13363 staphylococ
27	63.6	258	ETXD_STAU	P20723 staphylococ
28	63.6	258	ETXD_STAU	P20723 staphylococ
29	63.6	276	PYGL_SYNEL	P50039 synecococc
30	63.6	283	YJ81_YEAS	P47148 saccharomyc
31	63.6	351	PEXE_PICAN	P78723 pichia angu
32	63.6	375	NP4_HUMAN	P09733 homo sapien
33	63.6	385	TGT_ZYMO	P28720 zymomonas m

34	21	63.6	461	1	DISK_MOUSE	O61466 mus musculu
35	21	63.6	480	1	SYL_MOUSE	Q92517 mycobacteri
36	21	63.6	554	1	NBLA_MOUSE	P52963 mus musculu
37	21	63.6	684	1	FLID_HELPJ	Q92191 helicobacte
38	21	63.6	684	1	FLID_HELPJ	P96786 helicobacte
39	21	63.6	727	1	GPKM_HUMAN	P43304 homo sapien
40	21	63.6	727	1	GPKM_MOUSE	Q64521 mus musculu
41	21	63.6	899	1	ANDR_MOUSE	P19091 mus musculu
42	21	63.6	904	1	NIAL_TOBAC	P16105 nicotiana t
43	21	63.6	904	1	NIAL_TOBAC	P08509 nicotiana t
44	21	63.6	1184	1	XMS2_DROME	Q93399 drosophila
45	21	63.6	1305	1	GAK_RAT	P97874 ratius norv

#### ALIGNMENTS

```

RESULT 1
ID      RRP_P13H4      STANDARD;      PRT;      604 AA.
AC      P06162;
DT      01-JAN-1988 (Rel. 06, Last sequence update)
DT      01-AUG-1988 (Rel. 08, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      RNA polymerase alpha subunit (EC 2.7.7.48) (Nucleocapsid
DE      phosphoprotein).
GN      P.
OS      Human parainfluenza 3 virus (strain NIH 47885).
OC      Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC      Paramyxoviridae; Paramyxovirinae; Respirivirus.
OX      NCBI_TaxID=11217;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=86291173; PubMed=3016995;
RA      Luk D., Sanchez A., Banerjee A.K.;
RT      "Messenger RNA encoding the phosphoprotein (P) gene of human
RT      parainfluenza virus 3 is bicistronic.";
RL      Virology 153:318-325(1986).
RN      [2]
RP      REVISIONS.
RA      Luk D., Sanchez A., Banerjee A.K.;
RL      Virology 156:193-194(1987).
RN      [3]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=87044104; PubMed=3022477;
RA      Galinski M.S., Mink M.A., Lambert D.M., Wechsler S.L., Pons W.M.;
RT      "Molecular cloning and sequence analysis of the human parainfluenza 3
RT      virus mRNA encoding the P and C proteins.";
RL      Virology 155:46-60(1986).
RN      [4]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=87085486; PubMed=3025346;
RA      Spriggs M.K., Collins P.L.;
RT      "Sequence analysis of the P and C protein genes of human
RT      parainfluenza virus type 3: patterns of amino acid sequence homology
RT      among paramyxovirus proteins.";
RL      J. Gen. Virol. 67:2705-2719(1986).
CC      -1- FUNCTION: THIS PROTEIN IS PROBABLY A COMPONENT OF THE ACTIVE
CC      POLYMERASE. IT MAY FUNCTION IN TEMPLATE BINDING.
CC      -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC      (RNA)(N).
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
DR      EMBL: M14890; AAA46866.1; ALT_SEQ.
DR      EMBL: X04721; CAA28429.1; -.
DR      EMBL: M14932; AAA66818.1; -.

```

```

DR EMBL: D10029; BA00921.1; ALT_SEQ.
DR PIR: A27010; RRN2P3.
DR PIR: A24189; RRN2P4.
DR PIR: A26896; RRN2P5.
DR InterPro: IPR002693; Paramyxo_P.
DR Pfam: PF01806; Paramyxo_P.1.
KW Transferrase; RNA-directed RNA polymerase; Nucleocapsid;
KW Phosphorylation.
FT CONFLICT 198 198 P -> N (IN REF. 4).
FT CONFLICT 243 243 MISSING (IN REF. 3 AND 4).
FT CONFLICT 244 244 MISSING (IN REF. 4).
FT CONFLICT 316 316 O -> R (IN REF. 3).
FT CONFLICT 584 584 C -> S (IN REF. 4).
SQ SEQUENCE 604 AA; 67661 MW; D2E0CEFA7027D0BD CRC64;

Query Match 72.7%; Score 24; DB 1; Length 604;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12
Db 226 KXSSTHQEDD 235

RESULT 2
ETC2_STAAU STANDARD; PRT; 266 AA.
ID ETC2_STAAU
AC P01553;
DT 21-JUN-1986 (Rel. 01, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Enterotoxin type C-1 precursor (SEC1).
GN EMTCL.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_Taxid=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88038352; PubMed=2823067;
RA Bohach G.A., Schlievert P.M.;
RT "Nucleotide sequence of the staphylococcal enterotoxin C1 gene and
RT relatedness to other pyrogenic toxins.";
RL Mol. Gen. Genet. 209:15-20(1987).
RN [2]
RP SEQUENCE OF 28-266.
RX MEDLINE=83213327; PubMed=6189824;
RA Schmidt J.J., Spero L.;
RT "The complete amino acid sequence of staphylococcal enterotoxin C1.";
RL J. Biol. Chem. 258:6300-6306(1983).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION
CC STAPHYLOCOCCAL FOOD POISONING SYNDROME.
CC -1- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
CC EMBL: X05815; CAA29260.1; -.
DR PIR: A01816; ENSAC1.
DR PIR: S06356; S06356.
DR HSSP: P34071; 1SE2.
DR InterPro: IPR001961; Staph/Strep_toxin.
DR Pfam: PF01123; Staph_strep_toxin; 1.
DR Pfam: PF02876; Staph_strep_toxin; 1.
DR PRINTS: PR00279; BACTRTOXIN.
DR PROSITE: PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE: PS00278; STAPH_STREP_TOXIN_2; 1.

```

```

KW Enterotoxin; Toxin; Signal; Superantigen.
FT SIGNAL 1 27
FT CHAIN 28 266 ENTEROTOXIN TYPE C-1.
FT DISULFID 120 137
FT CONFLICT 177 177 D -> N (IN REF. 2).
SQ SEQUENCE 266 AA; 30546 MW; 3A7AB59A898653B CRC64;

Query Match 69.7%; Score 23; DB 1; Length 266;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12
Db 178 KXSVTAQELD 187

RESULT 3
ETC2_STAAU STANDARD; PRT; 266 AA.
ID ETC2_STAAU
AC P34071;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Enterotoxin type C-2 precursor (SEC2).
GN EMTCL.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_Taxid=1280;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-66.
RX MEDLINE=89277549; PubMed=2543637;
RA Bohach G.A., Schlievert P.M.;
RT "Conservation of the biologically active portions of staphylococcal
RT enterotoxins C1 and C2.";
RL Infect. Immun. 57:2249-2252(1989).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=96027099; PubMed=7582894;
RA Papageorgiou A.C., Acharya K.R., Shapiro R., Passalacqua E.F.,
RA Brehm R.D., Tranter H.S.;
RT "Crystal structure of the superantigen enterotoxin C2 from
RT Staphylococcus aureus reveals a zinc-binding site.";
RL Structure 3:769-779(1995).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RX MEDLINE=96022987; PubMed=7552730;
RA Swaminathan S., Furey W.F. Jr., Pletcher J., Sax M.;
RT "Residues defining V beta specificity in staphylococcal
RT enterotoxins.";
RL Nat. Struct. Biol. 2:680-686(1995).
RN [4]
RP COMPARISON OF STRUCTURE OF SEA AND SEC2.
RX MEDLINE=97343373; PubMed=9191070;
RA Schad E.M., Papageorgiou A.C., Svensson L.A., Acharya K.R.;
RT "A structural and functional comparison of staphylococcal
RT enterotoxins A and C2 reveals remarkable similarity and
RT dissimilarity.";
RL J. Mol. Biol. 269:270-280(1997).
CC -1- CORONATOR: Binds 1 zinc ion per subunit. The zinc ion is necessary
CC for the toxin interaction with MHC class II.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION
CC STAPHYLOCOCCAL FOOD POISONING SYNDROME.
CC -1- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN
CC FAMILY.
CC -----
DR PIR: A60114; A60114.
DR PDB: 1STE; 23-DEC-96.
DR PDB: 1SE2; 08-MAR-96.
DR InterPro: IPR001961; Staph/Strep_toxin.
DR Pfam: PF01123; Staph_strep_toxin; 1.
DR Pfam: PF02876; Staph_strep_toxin; 1.
DR PRINTS: PR00279; BACTRTOXIN.
DR PROSITE: PS00277; STAPH_STREP_TOXIN_1; 1.

```

DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.  
 KW Enterotoxin; Toxin; Signal; Superantigen; Metal-binding; Zinc;  
 3D-structure.  
 FT SIGNAL 1 27  
 FT CHAIN 28 266 ENTEROTOXIN TYPE C-2.  
 FT DISULFID 120 137  
 FT METAL 36 36 ZINC.  
 FT METAL 110 110 ZINC.  
 FT METAL 145 145 ZINC.  
 FT METAL 149 149 ZINC.  
 SQ SEQUENCE 266 AA; 30604 MW; 8407FBI8536FAC08 CRC64;  
 Query Match 69.7%; Score 23; DB 1; Length 266;  
 Best Local Similarity 50.0%; Pred. No. 20;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 3 KXXXTXQEXD 12  
 DB 178 KRSVTAQELD 187  
 RESULT 4  
 ETC3\_STAM STANDARD; PRT; 266 AA.  
 ID ETC3\_STAM  
 AC P2313;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Enterotoxin type C-3 precursor (SEC3).  
 GN EMT3 OR SAV2009 OR SA1817.  
 OS Staphylococcus aureus (strain N315) / ATCC 700699)  
 OS Staphylococcus aureus (strain N315), and  
 OS Staphylococcus aureus.  
 OC Bacteria: Firmicutes: Bacillales; Staphylococcus.  
 OX NCBI\_TaxID=158878, 158879, 1280;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-MU50 / ATCC 700699, and N315;  
 RX MEDLINE-21311952; PubMed-11418146;  
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,  
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,  
 RA Kanemori M., Matsunami H., Maruyama A., Murakami H., Hosoyma A.,  
 RA Mizutani-Uji Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,  
 RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,  
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,  
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,  
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus  
 aureus."  
 RL Lancet 357:1225-1240(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-90220508; PubMed-2325627;  
 RT "Nucleotide sequence of the staphylococcal enterotoxin C3 gene:  
 sequence comparison of all three type C staphylococcal  
 enterotoxins."  
 RL Mol. Gen. Genet. 220:329-333(1990).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS) OF COMPLEX WITH TCR.  
 RX MEDLINE-97064178; PubMed-8906797;  
 RA Fields B.A., Malchiodi E.L., Li H., Yern X., Stauffacher C.V.,  
 RA Schlievert P.M., Karjalainen K., Maritzza R.A.,  
 RT "Crystal structure of a T-cell receptor beta-chain complexed with a  
 superantigen."  
 RL Nature 384:188-192(1996).  
 RN [4]  
 RP SUBCELLULAR LOCATION: Secreted.  
 CC -1- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION  
 STAPHYLOCOCCAL FOOD POISONING SYNDROME.  
 CC -1- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN  
 FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>  
 CC or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch)).  
 CC -----  
 CC EMBL: AP003364; BAB58171.1; -;  
 CC EMBL: AP003135; BAB43097.1; -;  
 CC EMBL: X51661; CA35972.1; -;  
 CC PIR: S11885; S11885.  
 CC PDB: 1JCK; 12-NOV-97.  
 CC InterPro: IPR001961; Staph/Strep\_toxin.  
 CC Pfam: PF01123; Staph\_Strep\_toxin; 1.  
 CC Pfam: PF02876; Staph\_Strep\_tox\_C; 1.  
 CC PRINTS: PR00279; BACTRLOXIN.  
 DR PROSITE: PS00277; STAPH\_STREP\_TOXIN\_1; 1.  
 DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.  
 KW Enterotoxin; Toxin; Signal; Superantigen; 3D-structure;  
 Complete proteome.  
 FT SIGNAL 1 27  
 FT CHAIN 28 266 ENTEROTOXIN TYPE C-3.  
 FT DISULFID 120 137  
 SQ SEQUENCE 266 AA; 30671 MW; 5ED8A32D1FFCA59 CRC64;  
 QY 3 KXXXTXQEXD 12  
 DB 178 KRSVTAQELD 187  
 RESULT 5  
 RS1\_RHIME STANDARD; PRT; 568 AA.  
 ID RS1\_RHIME  
 AC P14129;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 30-JAN-2000 (Rel. 39, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 30S ribosomal protein S1.  
 GN RPSA OR R00255 OR SMC00335.  
 OS Rhizobium meliloti (Sinorhizobium meliloti).  
 OC Bacteria: Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae: Sinorhizobium.  
 OX NCBI\_TaxID=382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-RCR2011 / SU47;  
 RX MEDLINE-88217521; PubMed-3368316;  
 RA Schindler J., Thamm S., Lutz R., Hussain A., Falst G.,  
 RA Dobrinski B.,  
 RT "Cloning and characterization of a gene from Rhizobium meliloti 2011  
 coding for ribosomal protein S1."  
 RL Nucleic Acids Res. 16:3075-3089(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-1021;  
 RX MEDLINE-21396507; PubMed-11481430;  
 RA Capela D., Barloy-Hubier F., Gouzy J., Bothe G., Ampe F., Batut J.,  
 RA Boissard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
 RA Godt T., Goffeau A., Kahn D., Kiss E., Leleure V., Masny D.,  
 RA Pohl T., Portetelle D., Puehler A., Purnelle B., Rameberger U.,  
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Gallibert F.,  
 RT "Analysis of the chromosome sequence of the legume symbiont  
 Sinorhizobium meliloti strain 1021."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
 CC -1- FUNCTION: BINDS MRNA; THUS FACILITATING RECOGNITION OF THE  
 INITIATION POINT. IT IS NEEDED TO TRANSLATE MRNA WITH A SHORT  
 SHINE-DALGARNO (SD) PURINE-RICH SEQUENCE (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE S1P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -1- SIMILARITY: CONTAINS 6 S1 MOTIF DOMAINS.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb-sdb.ch](mailto:license@sdb-sdb.ch)).

DR EMBL; X07528; CAA30404.1; -  
 DR EMBL; AL591783; CAC41692.1; -  
 DR PIR; S01055; R3ZRL.  
 DR HSP; P05055; ISRO.  
 DR InterPro; IPR000110; Ribosomal\_S1.  
 DR InterPro; IPR003029; S1.  
 DR Pfam; PF00575; S1; 6.  
 DR PRINTS; PR00681; RIBOSOMAL\_S1.  
 DR SMART; SM00316; S1; 6.  
 DR TIGRfams; TIGR00717; rpsa; 1.  
 DR PROSITE; PS50126; S1; 6.  
 KW Ribosomal protein; Repeat; RNA-binding; Complete proteome.  
 FT DOMAIN 27 93 S1 MOTIF 1.  
 FT DOMAIN 111 177 S1 MOTIF 2.  
 FT DOMAIN 198 266 S1 MOTIF 3.  
 FT DOMAIN 283 353 S1 MOTIF 4.  
 FT DOMAIN 370 440 S1 MOTIF 5.  
 FT DOMAIN 459 530 S1 MOTIF 6.  
 FT CONFLICT 88 A -> R (IN REF. 1).  
 FT CONFLICT 153 162 LMHPQPEI -> ADAQPALRN (IN REF. 1).  
 FT CONFLICT 235 235 N -> K (IN REF. 1).  
 FT CONFLICT 241 241 L -> O (IN REF. 1).  
 FT CONFLICT 551 A -> R (IN REF. 1).  
 SQ SEQUENCE 568 AA; 62640 MW; AB0858204273A7B8 CRC64;

Query Match 69.7%; Score 23; DB 1; Length 568;  
 Best Local Similarity 50.0%; Pred. No. 45;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
 | | | | |  
 DB 326 KILTSQEVN 335

RESULT 6  
 ID ESRL\_XENLA STANDARD; PRT; 586 AA.  
 AC P81559;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Estrogen receptor (ER) (Estradiol receptor) (ER-alpha).  
 GN ESR1 OR NR3A1 OR ESR.  
 OS Xenopus laevis (African clawed frog).  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;  
 CC Xenopodidae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90331927; PubMed=1274894;  
 RA Weller I.J., Lew D., Shapiro D.J.;  
 RT "The Xenopus laevis estrogen receptor: sequence homology with human  
 RT and avian receptors and identification of multiple estrogen receptor  
 RT messenger ribonucleic acids.";  
 RL Mol. Endocrinol. 1:355-362(1987).  
 CC -1- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED  
 CC IN THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT  
 CC CELLULAR PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.  
 CC -1- SUBUNIT: BINDS DNA AS A HOMODIMER. CAN FORM A HETERODIMER WITH ER-  
 CC BETA (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,  
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.

## NR3 SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb-sdb.ch](mailto:license@sdb-sdb.ch)).

DR EMBL; L20735; -; NOT\_ANNOTATED\_CDS.  
 DR HSP; P03372; HRCQ.  
 DR TRANSFAC; T00263; -  
 DR InterPro; IPR000536; Hormone\_rec\_11g.  
 DR InterPro; IPR001292; Oestrgn\_receptor.  
 DR InterPro; IPR001723; Strhnm\_receptor.  
 DR InterPro; IPR001628; ZnF\_Casteroid.  
 DR Pfam; PF00104; hormone\_rec; 1.  
 DR Pfam; PF00105; zf-C4; 1.  
 DR Pfam; PF02159; Oest\_rec; 1.  
 DR PRINTS; PR00398; STRDHOMONER.  
 DR PRODOM; PD000035; ZnF\_Casteroid; 1.  
 DR SMART; SM00430; HOL1; 1.  
 DR SMART; SM00399; ZnF\_C4; 1.  
 DR PROSITE; PS00331; NUCLEAR\_RECEPTOR; 1.  
 KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;  
 FT ZINC-FINGER; Steroid-binding.  
 FT DOMAIN 1 179 MODULATING.  
 FT DNA\_BIND 180 245 NUCLEAR RECEPTOR-TYPE.  
 FT ZN\_FING 180 200 C4-TYPE.  
 FT ZN\_FING 216 240 C4-TYPE.  
 FT DOMAIN 246 302 HINGE.  
 FT DOMAIN 303 586 STEROID-BINDING.  
 SQ SEQUENCE 586 AA; 66080 MW; 0EDC77EBD06F08BF CRC64;

Query Match 69.7%; Score 23; DB 1; Length 586;  
 Best Local Similarity 50.0%; Pred. No. 46;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
 | | | | |  
 DB 545 KDKTTQEDN 554

RESULT 7  
 ID REEL\_ECOLI STANDARD; PRT; 251 AA.  
 AC P03856; P08969;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Replication initiation protein (RIP protein) (protein E) (F4 protein).  
 GN REP OR REP OR E.  
 OS Escherichia coli.  
 OX Escherichia coli.  
 OX Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OX Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=82051318; PubMed=7028572;  
 RA Murotsu T., Matsubara K., Sugisaki H., Takamami M.;  
 RT "Nine unique repeating sequences in a region essential for  
 RT replication and incompatibility of the mini-F plasmid.";  
 RL Gene 15:257-271(1981).  
 CC [2]  
 RN SEQUENCE FROM N.A.  
 RP Eichenlaub R.;  
 RA "F plasmid DNA complete mini-F region (F coordinates 40,301F to  
 RT 49,869F)." ;  
 RL Submitted (Aug-1986) to the EMBL/Genbank/DBJ databases.  
 RN [3]

RP SEQUENCE FROM N.A.  
RC STRAIN-K12 / CR63;  
RA Shimizu H., Saitoh Y., Suda Y., Uehara K., Sampel G., Mizobuchi K.;  
RT "Complete nucleotide sequence of the F plasmid: its implications for  
RT organization and diversification of F plasmid genomes.";  
RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE OF 211-251 FROM N.A.  
RX MEDLINE-8123454; PubMed-7018695;  
RA Tolun A., Helinski D.R.;  
RT "Direct repeats of the F plasmid *incC* region express F  
RT incompatibility.";  
RL Cell 24:687-694(1981).  
RN [5]  
RP SEQUENCE OF 1-17 FROM N.A.  
RX MEDLINE-86174346; PubMed-3007930;  
RA Disque-Kochem C., Seidel U., Helsenberg M., Eichenlaub R.;  
RT "The repeated sequences (*incB*) preceding the protein E gene of  
RT plasmid *minI-F* are essential for replication.";  
RL Mol. Genet. 202:132-135(1986).  
RN [6]  
RP SEQUENCE OF 209-251 FROM N.A.  
RC STRAIN-K12;  
RA MEDLINE-87141188; PubMed-3029390;  
RX Mori H., Kondo A., Oshima A., Ogura T., Hiraga S.;  
RT "Structure and function of the F plasmid genes essential for  
RT partitioning.";  
RL J. Mol. Biol. 192:1-15(1986).  
RN [7]  
RP SIMILARITY: REGIONS OF SIMILARITY WITH REPA PROTEIN OF THE P1  
PLASMID, BUT NOT WITH REPLICATION PROTEINS OF OTHER PLASMIDS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: J01724; AAA91062.1; -  
DR EMBL: M12987; AAA24901.1; -  
DR EMBL: AP001918; BA97915.1; -  
DR EMBL: X03410; CAA27146.1; -  
DR EMBL: X04619; CAA28294.1; -  
DR PIR: A04485; QOECF.  
DR PIR: D25783; D25783.  
DR Ecogene: EG40058; repe.  
DR InterPro: IPR000525; Initr\_Repr.  
DR Pfam: PF01051; RepB\_protein; 1.  
DR Plasmid; DNA replication; Plasmid copy control; Complete proteome.  
SQ SEQUENCE 251 AA; 29358 MW; EA2BC265D80158B2 CRC64;  
OY 3 KXXXTXQEXD 12  
DB 48 KSDGTQEHND 57  
Query Match 66.7%; Score 22; DB 1; Length 251;  
Best Local Similarity 50.0%; Pred. No. 36;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OX Streptococcus.  
NCBI\_TaxID=1314, 186103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-86166804; PubMed-3514452;  
RA Weeks C.R., Ferretti J.J.;  
RT "Nucleotide sequence of the type A streptococcal exotoxin  
RT (erythrogenic toxin) gene from Streptococcus pyogenes bacteriophage  
RT T12.";  
RL Infect. Immun. 52:144-150(1986).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-86284313; PubMed-3526093;  
RA Johnson L.P., L'Italian J.J., Schlievert P.M.;  
RT "Streptococcal pyrogenic exotoxin type A (scarlet fever toxin) is  
RT related to Staphylococcus aureus enterotoxin B.";  
RL Mol. Genet. 203:354-356(1986).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-MGAS8232 / Serotype M18;  
RX MEDLINE-21927593; PubMed-11917108;  
RA Smoot J.C., Barbien K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,  
RA Sylvia G.L., Sturdevant D.E., Ricklefs S.M., Porcilla S.F.,  
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,  
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;  
RT "Genome sequence and comparative microarray analysis of serotype M18  
RT group A Streptococcus strains associated with acute rheumatic fever  
RT outbreaks.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (2.57 ANGSTROMS).  
RX MEDLINE-99094887; PubMed-9878045;  
RA Papageorgiou A.C., Collins C.M., Gutman D.M., Kline J.B.,  
RA O'Brien S.M., Tranter H.S., Acharya K.R.;  
RT "Structural basis for the recognition of superantigen streptococcal  
RT pyrogenic exotoxin A (SpeA) by MHC class II molecules and T-cell  
RT receptors.";  
RL EMBO J. 18:9-21(1999).  
RN [5]  
RP SUBUNIT: Binds to major histocompatibility complex class II beta  
chain.  
CC -----  
CC -1- DISEASE: THE STREPTOCOCCAL PYROGENIC TOXINS A, B, AND C ARE  
CC THE CAUSATIVE AGENTS OF THE SYMPTOMS ASSOCIATED WITH SCARLET  
CC FEVER, HAVE BEEN ASSOCIATED WITH STREPTOCOCCAL TOXIC SHOCK-LIKE  
CC DISEASE AND MAY PLAY A ROLE IN THE EARLY EVENTS OF RHEUMATIC  
CC FEVER.  
CC -----  
CC -1- MISCELLANEOUS: THIS TOXIN SEEMS TO BE CODED BY BACTERIOPHAGE T12.  
CC -1- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN  
CC FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: U40453; AAC48868.1; -  
DR EMBL: X03929; CAA27568.1; -  
DR EMBL: AE009982; AAL97141.1; -  
DR PIR: A26152; A26152.  
DR PIR: S29659; S29659.  
DR PDB: 1B12; 24-NOV-99.  
DR InterPro: IPR001961; Staph/Strep\_toxin.  
DR Pfam: PF01123; Staph-Strep\_toxin; 1.  
DR Pfam: PF02876; Staph-Strep\_tox\_C; 1.  
DR PRINTS: PR00279; BACTRLTOXIN.  
DR PROSITE: PS00277; STAPH\_STREP\_TOXIN\_1; 1.  
DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.  
KW Toxin; Signal; 3D-structure.  
FT STGNAL 1 30  
FT CHAIN 31 251 EXOTOXIN TYPE A.

```

FT DISULFID 117 128 K -> E (IN REF. 2).
FT CONFLICT 6 6 VT -> MK (IN REF. 2).
FT CONFLICT 17 18 SOEFAODPD -> LFGICSTRPK (IN REF. 2).
FT CONFLICT 25 35 H -> O (IN REF. 2).
FT CONFLICT 40 40 S -> N (IN REF. 2).
FT CONFLICT 43 43 NLOMIFLEBDP -> TFKIIFEMRYTL (IN
FT CONFLICT 47 59 REF. 2).
FT CONFLICT 129 129 I -> L (IN REF. 2).
FT CONFLICT 165 178 TNKRMVTAQELDYK -> QIRNGMCSRSYR (IN
FT CONFLICT 165 178 REF. 2).
SQ SEQUENCE 251 AA; 29246 MM; 54001FEACCCAFCC3 CRC64;

Query Match 66.7%; Score 22; DB 1; Length 251;
Best Local Similarity 50.0%; Pred. No. 36;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXOEXD 12
Db 167 KKVTAQELD 176

RESULT 9
ETXB_STAAU STANDARD; PRT; 266 AA.
AC P01532;
DT 21-JUL-1986 (Rel. 01, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Enterotoxin type B precursor (SEB).
GN ENTB.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID:1280;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE-86168029; PubMed-3957869;
RA Jones C.L., Khan S.A.;
RT "Nucleotide sequence of the enterotoxin B gene from Staphylococcus
RT aureus."
RL J. Bacteriol. 166:29-33(1986).
[2]
RP SEQUENCE OF 40-91 FROM N.A.
RX MEDLINE-85298255; PubMed-3898073;
RA Raneelli D.M., Jones C.L., Johns M.B., Mussey G.J., Khan S.A.;
RT "Molecular cloning of staphylococcal enterotoxin B gene in
RT Escherichia coli and Staphylococcus aureus."
RL Proc. Natl. Acad. Sci. U.S.A. 82:5850-5854(1985).
[3]
RP SEQUENCE OF 28-266 (S-6).
RX MEDLINE-71007902; PubMed-5470821;
RA Huang I.-Y., Bergdoll M.S.;
RT "The primary structure of staphylococcal enterotoxin B. 3. The
RT cytosine bromide peptides of reduced and aminoethylated enterotoxin
RT B, and the complete amino acid sequence."
RL J. Biol. Chem. 245:3518-3525(1970).
[4]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE-93063291; PubMed-1436058;
RA Swaminathan S., Furey W.F., Jr., Pletcher J., Sax M.;
RT "Crystal structure of staphylococcal enterotoxin B, a superantigen."
RL Nature 359:801-806(1992).
[5]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF COMPLEX WITH MHC II.
RX MEDLINE-94203283; PubMed-8152483;
RA Janderzky T.S., Brown J.H., Gorga J.C., Stern L.J., Urban R.G.;
RT "Three-dimensional structure of a human class II histocompatibility
RT molecule complexed with superantigen."
RL Nature 368:711-718(1994).
[6]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF COMPLEX WITH TCR.
RX MEDLINE-99096298; PubMed-9881971;

```

```

RA Li H., Llera A., Tsuchiya D., Leder L., Ysern X., Schlievert P.M.,
RA Karjalainen K., Marizuka R.A.;
RT "Three-dimensional structure of the complex between a T cell receptor
RT beta chain and the superantigen staphylococcal enterotoxin B."
RL Immunity 9:807-816(1998).
[7]
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RX MEDLINE-98181012; PubMed-9514739;
RA Papageorgiou A.C., Tranter H.S., Acharya K.R.;
RT "Crystal structure of microbial superantigen staphylococcal
RT enterotoxin B at 1.5-A resolution: implications for superantigen
RT recognition by MHC class II molecules and T-cell receptors."
RL J. Mol. Biol. 277:61-79(1998).
CC -1 SUBCELLULAR LOCATION: Secreted.
CC -1 DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION
CC STAPHYLOCOCCAL FOOD POISONING SYNDROME.
CC -1 SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M11118; AAA89550.1; -.
DR PIR: A01815; ENSGAB6.
DR PIR: S27360; S27360.
DR PDB: 1SEB; 20-JUN-96.
DR PDB: 2SEB; 28-JAN-98.
DR PDB: 3SEB; 27-MAY-98.
DR PDB: 1SE3; 16-JUN-97.
DR PDB: 1SE4; 15-OCT-97.
DR PDB: 1SEB; 04-MAR-99.
DR InterPro: IPR001961; Scap/Strep_toxin.
DR Pfam: PF01123; Strep_strep_toxin; 1.
DR Pfam: PF02876; Strep_strep_toxin; 1.
DR PRINTS: PR00279; BACTRTOXIN.
DR PROSITE: PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE: PS00278; STAPH_STREP_TOXIN_2; 1.
DR KAT Enterotoxin; Toxin; Signal; Superantigen; 3D-structure.
FT SIGNAL 1 27
FT CHAIN 28 266 ENTEROTOXIN TYPE B.
FT DISULFID 120 140
FT CONFLICT 56 58
FT CONFLICT 69 77 DDN -> NND (IN REF. 3).
FT CONFLICT 118 118 DQIYFDLI -> NEFDLIYL (IN REF. 3).
FT CONFLICT 128 130 MISSING (IN REF. 3).
FT CONFLICT 133 135 DIN -> NID (IN REF. 3).
FT CONFLICT 149 150 QTD -> ENT (IN REF. 3).
FT CONFLICT 156 156 NG -> GN (IN REF. 3).
FT CONFLICT 185 186 Y -> YV (IN REF. 3).
FT CONFLICT 185 186 OE -> EO (IN REF. 3).
FT CONFLICT 233 233 D -> N (IN REF. 3).
FT CONFLICT 246 247 DN -> ND (IN REF. 3).
SQ SEQUENCE 266 AA; 31436 MM; B6D417F61CF01B80 CRC64;

Query Match 66.7%; Score 22; DB 1; Length 266;
Best Local Similarity 50.0%; Pred. No. 39;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXOEXD 12
Db 179 KKVTAQELD 188

RESULT 10
G3P3_YEAST STANDARD; PRT; 331 AA.
AC P00359;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)

```

DT 15-JUN-2002 (rel. 41, last annotation update)  
 DE glyceraldehyde 3-phosphate dehydrogenase 3 (EC 1.2.1.12) (GAPDH 3).  
 GN IDH3 OR GPD3 OR YGR192C OR G7576.  
 OS Saccharomyces cerevisiae (Baker's Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 ON NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=80027306; PubMed=385592;  
 RA Holland J.P., Holland M.J.,  
 RT "The primary structure of a glyceraldehyde-3-phosphate dehydrogenase  
 gene from Saccharomyces cerevisiae."  
 RL J. Biol. Chem. 254:9839-9845(1979).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=S288C;  
 RC MEDLINE=95373283; PubMed=7645350;  
 RA Arroyo J., Garcia-Gonzalez M., Garcia-Saez M.I., Sanchez M.,  
 RM Nombela C.;  
 RT "The complete sequence of a 9037 bp DNA fragment of the right arm of  
 Saccharomyces cerevisiae chromosome VII."  
 RL Yeast 11:587-591(1995).  
 RN [3]  
 RP SEQUENCE OF 46-57.  
 RX STRAIN=S288C;  
 RC MEDLINE=95203288; PubMed=7895733;  
 RA Garrels J.I., Fletcher B., Kobayashi R., Latter G.I., Schwender B.,  
 RM Volpe T., Warner J.R., McLaughlin C.S.;  
 RT "Protein identifications for a Saccharomyces cerevisiae protein  
 database."  
 RL Electrophoresis 15:1466-1486(1994).  
 RN [4]  
 RP PARTIAL SEQUENCE.  
 RX STRAIN=ATCC 38531 / Y41, and SKO2N;  
 RC MEDLINE=95255188; PubMed=7737086;  
 RA Norbeck J., Blomberg A.;  
 RT "Gene linkage of two-dimensional polyacrylamide gel electrophoresis  
 resolved proteins from isogene families in Saccharomyces cerevisiae  
 by microsequencing of in-gel trypsin generated peptides."  
 RL Electrophoresis 16:149-156(1995).  
 CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +  
 NAD(+) -> 3-phospho-D-glyceroyl phosphate + NADH.  
 CC -1- PATHWAY: Second phase of glycolysis; first step.  
 CC -1- SUBUNIT: HOMOTETRAMER.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- MISCELLANEOUS: THERE ARE THREE GENES FOR G3PDH IN YEAST.  
 CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE  
 DEHYDROGENASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: J01300; CAA24607.1; -;  
 DR EMBL: J01324; AAA88714.1; -;  
 DR EMBL: X82408; CAA57803.1; -;  
 DR EMBL: Z72977; CAA97218.1; -;  
 DR PIR: A00371; DEBYG2.  
 DR HSSP: P06977; LGAD.  
 DR SWISS-2DPAGE: P00359; YEAST.  
 DR COMPUYEST-2DPAGE: P00359; -;  
 DR SGD: S0003424; TDH3.  
 DR InterPro: IPR000173; GAP\_dhydrogenase.  
 DR Pfam: PF00044; gpdh.1.  
 DR Pfam: PF02800; gpdh.C.1.  
 DR PRINTS: PR00078; G3PDHGNASE.  
 DR PROSITE: PS00071; GAPDH; 1.  
 KW Glycolysis; Oxidoreductase; NAD; Multigene family.

FT INIT\_MET 0 0  
 FT BINDING 149 149 GLYCERALDEHYDE 3-PHOSPHATE.  
 FT ACT\_SITE 176 176 ACTIVATES THIOL GROUP DURING CATALYSIS.  
 FT CONFLICT 135 135 E -> V (IN REF. 1).  
 FT CONFLICT 247 247 N -> D (IN REF. 1).  
 FT CONFLICT 328 328 V -> I (IN REF. 1).  
 SQ SEQUENCE 331 AA; 35615 MW; CFE94A335C648B5 CRC64;  
 Query Match 66.7%; Score 22; DB 1; Length 331;  
 Best Local Similarity 50.0%; Pred. No. 49;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 3 KXXXTQXEND 12  
 Db 69 KRATYQERD 78  
 RESULT 11  
 DKK3\_CHICK STANDARD; PRT; 350 AA.  
 AC 090839;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE DICKKOPF related protein-3 precursor (Dkk-3) (DICKKOPF-3) (Lens fiber  
 DE protein CLFEFT4).  
 GN DKK3.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 ON NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lens fibers;  
 RX MEDLINE=96437509; PubMed=8840185;  
 RA Sawada K., Agata K., Eguchi G.;  
 RT "Characterization of terminally differentiated cell state by  
 RT categorizing cDNA clones derived from chicken lens fibers."  
 RL Int. J. Dev. Biol. 40:531-535(1996).  
 CC -1- FUNCTION: INHIBITOR OF WNT SIGNALING PATHWAY (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: secreted (Potential).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN EYE LENS.  
 CC -1- SIMILARITY: BELONGS TO THE DICKKOPF FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: D26311; BAA05373.1; -;  
 DR HSSP: P25687; IIMT.  
 KW Developmental Protein; Signal; Glycoprotein.  
 FT SIGNAL 1 29  
 FT CHAIN 30 350  
 FT DOMAIN 139 187  
 FT DOMAIN 200 277  
 FT CARBOHYD 88 88  
 FT CARBOHYD 98 98  
 FT CARBOHYD 113 113  
 FT CARBOHYD 196 196  
 FT CARBOHYD 282 282  
 SQ SEQUENCE 350 AA; 39208 MW; 57BE7ED850089DAE CRC64;  
 Query Match 66.7%; Score 22; DB 1; Length 350;  
 Best Local Similarity 50.0%; Pred. No. 52;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 3 KXXXTQXEND 12  
 I I I I I

Db 99 KTVOTHOEID 108

RESULT 12

YDGC\_SCHPO

ID YDGC\_SCHPO STANDARD: PRT: 356 AA.

AC 010498;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Hypothetical protein C26F1.12c in chromosome I.

SPAC26F1.12c.

OS Schizosaccharomyces pombe (fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetaceae; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

NCBI\_TaxID=4896;

[1]

SEQUENCE FROM N.A.

RC STRAIN-972;

RA MEDLINE-21848401; PubMed-11859360;

RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,

RA Brooks K., Brown S., Chillingworth T., Churcher C.M.,

RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,

RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,

RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,

RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,

RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,

RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,

RA Skelton J., Simmonds M., Squares R., Squares K., Stevens K.,

RA Taylor K., Taylor R.G., Tiley A., Walsh S.V., Warren T., Whitehead S.,

RA Woodward J., Volkart G., Aert R., Robben J., Grymponprez B.,

RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,

RA Gabriel C., Fuchs M., Fritz C., Holzer E., Moesli D., Hilbert H.,

RA Borzym K., Langer I., Beck A., Lehman H., Reinhardt R., Pohl T.M.,

RA Eger P., Zimmermann W., Wedler H., Wambolt R., Purnelle B.,

RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motlier S.,

RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,

RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,

RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,

RA Cerruti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,

RA Shpakovski G.V., Useary D., Barrett B.G., Nurse P.;

RA "The genome sequence of Schizosaccharomyces pombe.";

RL Nature 415:871-880(2002).

CC -1 SIMILARITY: TO YEAST HGHI.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@sib-sib.ch).

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

AC 092CA5;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein RP855.

RP855.

OS Rickettsia prowazekii.

OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;

OC Rickettsiaceae; Rickettsiae; Rickettsia.

NCBI\_TaxID=782;

[1]

SEQUENCE FROM N.A.

RC STRAIN-Madrid E;

RA MEDLINE-99039499; PubMed-9823893;

RA Anderson S.G.E., Zomrodipour A., Anderson J.O.,

RA Sichert-Bonten T., Alsmark U.C.M., Podowski R.M., Naeisund A.K.,

RA Eriksson A.-S., Winkler H.H., Kuriland C.G.;

RA "The genome sequence of Rickettsia prowazekii and the origin of

RA mitochondria.";

RL Nature 396:133-140(1998).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@sib-sib.ch).

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

AC 092CA5;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein RP855.

RP855.

OS Rickettsia prowazekii.

OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;

OC Rickettsiaceae; Rickettsiae; Rickettsia.

NCBI\_TaxID=782;

[1]

SEQUENCE FROM N.A.

RC STRAIN-Madrid E;

RA MEDLINE-99039499; PubMed-9823893;

RA Anderson S.G.E., Zomrodipour A., Anderson J.O.,

RA Sichert-Bonten T., Alsmark U.C.M., Podowski R.M., Naeisund A.K.,

RA Eriksson A.-S., Winkler H.H., Kuriland C.G.;

RA "The genome sequence of Rickettsia prowazekii and the origin of

RA mitochondria.";

RL Nature 396:133-140(1998).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@sib-sib.ch).

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC



KW Hypothetical protein.  
SQ SEQUENCE 712 AA; 80794 MW; 54B051ADPA547DE CRC64;

Query Match  
Best Local Similarity 50.0%; Score 22; DB 1; Length 712;  
Pred. No. 1.1e+02;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
DB 88 KKFITSQEDD 97

RESULT 15  
YE86\_SCHPO  
ID YE86\_SCHPO STANDARD; PRT; 886 AA.  
AC 014302;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Hypothetical protein C9G1.06c in chromosome I.  
GN SPAC9G1.06c.  
OS Schizosaccharomyces pombe (fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE-21848401; PubMed-11859360;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagers K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabbinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
RA Taylor K., Taylor R.C., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volkart G., Aert R., Robben J., Grymoprez B.,  
RA Wellens I., Vanstreels E., Aert R., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Medler H., Wandt R., Purnelle B.,  
RA Coffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,  
RT The genome sequence of Schizosaccharomyces pombe.  
RL Nature 415:871-880(2002).  
CC -1- SIMILARITY: TO YEAST YDL117W.  
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL; Z98763; CAB11490.1; -  
CC DR HSSP; P29355; ISEM.  
CC DR InterPro; IPR001452; SH3.  
CC DR InterPro; IPR002931; Trnsglucanase\_like.  
CC DR Pfam; PF00018; SH3; 1.  
CC DR ProDom; PD000066; SH3; 1.

DR SMART; SM00326; SH3; 1.  
DR SMART; SM00460; TGC; 1.  
DR PROSITE; PS50002; SH3; 1.  
KW Hypothetical protein: SH3 domain.  
FT DOMAIN 6 67 SH3.  
FT DOMAIN 607 610 POLY-SER.  
SQ SEQUENCE 886 AA; 98262 MW; D0C4879C1882E869 CRC64;

Query Match  
Best Local Similarity 40.0%; Score 22; DB 1; Length 886;  
Pred. No. 1.4e+02;  
Matches 4; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
DB 220 KSTTSKETD 229

Search completed: June 5, 2003, 15:56:34  
Job time: 7.81818 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: June 5, 2003, 15:32:58 ; Search time 34.9091 Seconds  
(without alignments)  
70.829 Million cell updates/sec

Title: US-09-150-947f-16  
Perfect score: 33  
Sequence: 1 XXXXXXTXQEXD 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database :

SPREMBL\_21: \*  
1: sp\_archaea: \*  
2: sp\_bacteria: \*  
3: sp\_fungi: \*  
4: sp\_human: \*  
5: sp\_invertebrate: \*  
6: sp\_mammal: \*  
7: sp\_mhc: \*  
8: sp\_organelle: \*  
9: sp\_phage: \*  
10: sp\_plant: \*  
11: sp\_rodent: \*  
12: sp\_virus: \*  
13: sp\_vertebrate: \*  
14: sp\_unclassified: \*  
15: sp\_virus: \*  
16: sp\_bacteriophage: \*  
17: sp\_archaeal: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	72.7	209	09R005	09R95 streptococ
2	24	72.7	209	09LAE1	09LAE1 streptococ
3	24	72.7	209	09LAE0	09LAE0 streptococ
4	24	72.7	209	09LAD8	09LAD8 streptococ
5	24	72.7	209	09LAD7	09LAD7 streptococ
6	24	72.7	209	09LAD6	09LAD6 streptococ
7	24	72.7	209	09LAD5	09LAD5 streptococ
8	24	72.7	209	09LAD4	09LAD4 streptococ
9	24	72.7	209	09LAD3	09LAD3 streptococ
10	24	72.7	209	09LAD2	09LAD2 streptococ
11	24	72.7	209	09LAD1	09LAD1 streptococ
12	24	72.7	209	09LAD0	09LAD0 streptococ
13	24	72.7	209	09LAC9	09LAC9 streptococ
14	24	72.7	209	09LAC8	09LAC8 streptococ
15	24	72.7	209	09LAC7	09LAC7 streptococ
16	24	72.7	209	09LAC7	09LAC7 streptococ

17	24	72.7	209	2	09LAC6	09LAC6 streptococ
18	24	72.7	209	2	09LAC5	09LAC5 streptococ
19	24	72.7	209	2	09LAC4	09LAC4 streptococ
20	24	72.7	209	2	09LAC3	09LAC3 streptococ
21	24	72.7	209	2	09LAC2	09LAC2 streptococ
22	24	72.7	209	2	09LAC1	09LAC1 streptococ
23	24	72.7	209	2	09LAC0	09LAC0 streptococ
24	24	72.7	209	2	09LAC0	09LAC0 streptococ
25	24	72.7	209	2	09LAC0	09LAC0 streptococ
26	24	72.7	209	2	09LAC0	09LAC0 streptococ
27	24	72.7	209	2	09LAC0	09LAC0 streptococ
28	24	72.7	209	2	09LAC0	09LAC0 streptococ
29	24	72.7	209	2	09LAC0	09LAC0 streptococ
30	24	72.7	209	2	09LAC0	09LAC0 streptococ
31	24	72.7	209	2	09LAC0	09LAC0 streptococ
32	24	72.7	209	2	09LAC0	09LAC0 streptococ
33	24	72.7	209	2	09LAC0	09LAC0 streptococ
34	24	72.7	209	2	09LAC0	09LAC0 streptococ
35	24	72.7	209	2	09LAC0	09LAC0 streptococ
36	24	72.7	209	2	09LAC0	09LAC0 streptococ
37	24	72.7	209	2	09LAC0	09LAC0 streptococ
38	24	72.7	209	2	09LAC0	09LAC0 streptococ
39	24	72.7	209	2	09LAC0	09LAC0 streptococ
40	24	72.7	209	2	09LAC0	09LAC0 streptococ
41	24	72.7	209	2	09LAC0	09LAC0 streptococ
42	24	72.7	209	2	09LAC0	09LAC0 streptococ
43	24	72.7	209	2	09LAC0	09LAC0 streptococ
44	24	72.7	209	2	09LAC0	09LAC0 streptococ
45	24	72.7	209	2	09LAC0	09LAC0 streptococ

## ALIGNMENTS

RESULT 1  
ID 09R005 PRELIMINARY; PRT; 209 AA.  
AC 09R005;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Mitogenic exotoxin Z 2 (Fragment).  
GN SMEZ-2.  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
OC Streptococcaceae; Streptococcus.  
OX NCBI\_TaxID-1314;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN-2035;  
RC MEDLINE-99093428; PubMed-9874566;  
RA Prof. T. Moffatt S.L., Berkahn C.J., Fraser J.D.;  
RT Identification and characterization of novel superantigens from  
RT Streptococcus pyogenes.  
RL J. Exp. Med. 189:89-102(1999).  
DR EMBL; AF086626; AAD52087.1; -.  
DR HSSP; P13163; 1SXT.  
DR InterPro; IPR001961; Staph/Strep toxin.  
DR Pfam; PF01123; Staph-Strep toxin; 1.  
DR PRINTS; PR00279; BACTRITOXIN.  
DR PROSITE; PS00278; STAPH\_STREP\_TOXIN\_2; 1.  
FT NON TER 1  
SQ SEQUENCE 209 AA; 24131 MW; 52BF7911BB100152 CRC64;

Query Match 3 XXXXTXQEXD 12  
Best Local Similarity 72.7%; Score 24; DB 2; Length 209;  
Matches 5; Conservative 0; Pred. No. 46;  
Mismatches 5; Indels 0; Gaps 0;

DB 125 KTTVTAQED 134

## RESULT 2

09LAE1 PRELIMINARY; PRT: 209 AA.  
ID 09LAE1;  
AC 09LAE1;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)  
DE Mitogenic exotoxin Z-3 (Fragment).  
GN SMEZ-3.  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
OC Streptococcaceae; Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-11681;  
RX MEDLINE-20273982; PubMed-10811869;  
RA Proft T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,  
Fraser J.D.;  
RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,  
Mosaic Structure, and Significant Antigenic Variation.";  
RL J. Exp. Med. 191:1765-1776(2000).  
DR EMBL: AF143653; AAF6654.1; -  
DR HSSP: P13163; 1SXT.  
DR InterPro: IPR001961; Stap/Strep\_toxin.  
DR Pfam: PF01123; Stap\_Strep\_toxin; 1.  
DR Pfam: PF02876; Stap\_Strep\_tox\_C; 1.  
DR PRINTS: PR00279; BACTRLTOXIN.  
DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.  
FT NON\_TER  
SQ SEQUENCE 209 AA; 24071 MW; F0ADFCDD1AA87271 CRC64;

Query Match 72.7%; Score 24; DB 2; Length 209;  
Best Local Similarity 50.0%; Pred. No. 46;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
DB 125 KTYVTAQED 134

## RESULT 3

09LAE0 PRELIMINARY; PRT: 209 AA.  
ID 09LAE0;  
AC 09LAE0;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)  
DE Mitogenic exotoxin Z-4 (Fragment).  
GN SMEZ-4.  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
OC Streptococcaceae; Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-9893;  
RX MEDLINE-20273982; PubMed-10811869;  
RA Proft T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,  
Fraser J.D.;  
RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,  
Mosaic Structure, and Significant Antigenic Variation.";  
RL J. Exp. Med. 191:1765-1776(2000).  
DR EMBL: AF143654; AAF6655.1; -  
DR HSSP: P13163; 1SXT.  
DR InterPro: IPR001961; Stap/Strep\_toxin.  
DR Pfam: PF01123; Stap\_Strep\_toxin; 1.  
DR Pfam: PF02876; Stap\_Strep\_tox\_C; 1.  
DR PRINTS: PR00279; BACTRLTOXIN.  
DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.  
FT NON\_TER  
SQ SEQUENCE 209 AA; 24108 MW; 67EC279BCC4A8247 CRC64;

Query Match 72.7%; Score 24; DB 2; Length 209;  
Best Local Similarity 50.0%; Pred. No. 46;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
DB 125 KTYVTAQED 134

## RESULT 4

09LAD9 PRELIMINARY; PRT: 209 AA.  
ID 09LAD9;  
AC 09LAD9;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)  
DE Mitogenic exotoxin Z-5 (Fragment).  
GN SMEZ-5.  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
OC Streptococcaceae; Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-11244;  
RX MEDLINE-20273982; PubMed-10811869;  
RA Proft T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,  
Fraser J.D.;  
RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,  
Mosaic Structure, and Significant Antigenic Variation.";  
RL J. Exp. Med. 191:1765-1776(2000).  
DR EMBL: AF143655; AAF6656.1; -  
DR HSSP: P13163; 1SXT.  
DR InterPro: IPR001961; Stap/Strep\_toxin.  
DR Pfam: PF01123; Stap\_Strep\_toxin; 1.  
DR Pfam: PF02876; Stap\_Strep\_tox\_C; 1.  
DR PRINTS: PR00279; BACTRLTOXIN.  
DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.  
FT NON\_TER  
SQ SEQUENCE 209 AA; 24079 MW; 24CA3885469CB9B CRC64;

Query Match 72.7%; Score 24; DB 2; Length 209;  
Best Local Similarity 50.0%; Pred. No. 46;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
DB 125 KTYVTAQED 134

## RESULT 5

09LAD8 PRELIMINARY; PRT: 209 AA.  
ID 09LAD8;  
AC 09LAD8;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)  
DE Mitogenic exotoxin Z-7 (Fragment).  
GN SMEZ-7.  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
OC Streptococcaceae; Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-11574;  
RX MEDLINE-20273982; PubMed-10811869;  
RA Proft T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,  
Fraser J.D.;  
RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,  
Mosaic Structure, and Significant Antigenic Variation.";  
RL J. Exp. Med. 191:1765-1776(2000).

DR EMBL: AF143657; AAF66657.1; -.  
 DR HSSP: P13163; ISXT.  
 DR InterPro: IPR001961; Staph/Strep\_toxin.  
 DR Pfam: PF01123; Staph\_strep\_toxin; 1.  
 DR Pfam: PF02876; Staph\_strep\_toxin; 1.  
 DR PRINTS: PR00279; BACTRTOXIN.  
 DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 209 AA; 24064 MW; 67E977CD5AA934F7 CRC64;

Query Match 72.7%; Score 24; DB 2; Length 209;  
 Best Local Similarity 50.0%; Pred. No. 46;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12  
 | | | |  
 Db 125 KTTVTAQED 134

RESULT 6  
 Q9LAD7 PRELIMINARY; PRT; 209 AA.  
 AC Q9LAD7;  
 DT 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
 DE Mitogenic exotoxin z-8 (Fragment).  
 GN SMEZ-8.  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
 OC Streptococcaceae; Streptococcus.  
 OX NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-1/5045;  
 RX MEDLINE=20273982; PubMed=10811869;  
 RA Proft T., Mofitt S.L., Weller K.D., Paterson A., Martin D.,  
 RA Fraser J.D.;  
 RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,  
 RT Mosaic Structure, and Significant Antigenic Variation."  
 RL J. Exp. Med. 191:1765-1776(2000).  
 DR EMBL: AF143658; AAF66658.1; -.  
 DR HSSP: P13163; ISXT.  
 DR InterPro: IPR001961; Staph/Strep\_toxin.  
 DR Pfam: PF01123; Staph\_strep\_toxin; 1.  
 DR Pfam: PF02876; Staph\_strep\_toxin; 1.  
 DR PRINTS: PR00279; BACTRTOXIN.  
 DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 209 AA; 24091 MW; 7680E9B35513DF28 CRC64;

Query Match 72.7%; Score 24; DB 2; Length 209;  
 Best Local Similarity 50.0%; Pred. No. 46;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12  
 | | | |  
 Db 125 KTTVTAQED 134

RESULT 7  
 Q9LAD6 PRELIMINARY; PRT; 209 AA.  
 AC Q9LAD6;  
 DT 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
 DE Mitogenic exotoxin z-9 (Fragment).  
 GN SMEZ-9.  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
 OC Streptococcaceae; Streptococcus.  
 OX NCBI\_TaxID=1314;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-11299;  
 RX MEDLINE=20273982; PubMed=10811869;  
 RA Proft T., Mofitt S.L., Weller K.D., Paterson A., Martin D.,  
 RA Fraser J.D.;  
 RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,  
 RT Mosaic Structure, and Significant Antigenic Variation."  
 RL J. Exp. Med. 191:1765-1776(2000).  
 DR EMBL: AF143659; AAF66659.1; -.  
 DR HSSP: P13163; ISXT.  
 DR InterPro: IPR001961; Staph/Strep\_toxin.  
 DR Pfam: PF01123; Staph\_strep\_toxin; 1.  
 DR Pfam: PF02876; Staph\_strep\_toxin; 1.  
 DR PRINTS: PR00279; BACTRTOXIN.  
 DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 209 AA; 24213 MW; B21587BA343B5DD5 CRC64;

Query Match 72.7%; Score 24; DB 2; Length 209;  
 Best Local Similarity 50.0%; Pred. No. 46;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12  
 | | | |  
 Db 125 KTTVTAQED 134

RESULT 8  
 Q9LAD5 PRELIMINARY; PRT; 209 AA.  
 AC Q9LAD5;  
 DT 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Mitogenic exotoxin z-10 (Fragment).  
 GN SMEZ-10.  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
 OC Streptococcaceae; Streptococcus.  
 OX NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-11580;  
 RX MEDLINE=20273982; PubMed=10811869;  
 RA Proft T., Mofitt S.L., Weller K.D., Paterson A., Martin D.,  
 RA Fraser J.D.;  
 RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,  
 RT Mosaic Structure, and Significant Antigenic Variation."  
 RL J. Exp. Med. 191:1765-1776(2000).  
 DR EMBL: AF143660; AAF66660.1; -.  
 DR HSSP: P13163; ISXT.  
 DR InterPro: IPR001961; Staph/Strep\_toxin.  
 DR Pfam: PF01123; Staph\_strep\_toxin; 1.  
 DR Pfam: PF02876; Staph\_strep\_toxin; 1.  
 DR PRINTS: PR00279; BACTRTOXIN.  
 DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 209 AA; 24102 MW; 8B4C853751A56D06 CRC64;

Query Match 72.7%; Score 24; DB 2; Length 209;  
 Best Local Similarity 50.0%; Pred. No. 46;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12  
 | | | |  
 Db 125 KTTVTAQED 134

RESULT 9  
 Q9LAD4 PRELIMINARY; PRT; 209 AA.  
 AC Q9LAD4;

DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Mitogenic exotoxin Z-11 (Fragment).  
 GN SMEZ-11.  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
 OC Streptococcaceae; Streptococcus.  
 OX NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=979;  
 RX MEDLINE=20273982; PubMed=10811869;  
 RA Prof. T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,  
 Fraser J.D.;  
 RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,  
 Mosaic Structure, and Significant Antigenic Variation.";  
 RL J. Exp. Med. 191:1765-1776(2000).  
 DR EMBL; AF143661; AAF66661.1; -.  
 DR HSP; P13163; ISXT.  
 DR InterPro: IPR001961; Strep/Strep-toxin.  
 DR Pfam: PF01123; Strep\_Strep-toxin; 1.  
 DR Pfam: PF02876; Strep\_Strep-toxin; 1.  
 DR PRINTS: PR00279; BACTRLTOXIN.  
 DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.  
 FT NON\_TER  
 SQ SEQUENCE 209 AA; 24194 MW; E12EF47B388D95DE CRC64;

Query Match 72.7%; Score 24; DB 2; Length 209;  
 Best Local Similarity 50.0%; Pred. No. 46;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12  
 DB 125 KTTVTAQAEID 134

RESULT 10  
 Q91AD3 PRELIMINARY; PRT; 209 AA.  
 AC Q91AD3;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Mitogenic exotoxin Z-12 (Fragment).  
 GN SMEZ-12.  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
 OC Streptococcaceae; Streptococcus.  
 OX NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=10303;  
 RX MEDLINE=20273982; PubMed=10811869;  
 RA Prof. T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,  
 Fraser J.D.;  
 RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,  
 Mosaic Structure, and Significant Antigenic Variation.";  
 RL J. Exp. Med. 191:1765-1776(2000).  
 DR EMBL; AF143662; AAF66662.1; -.  
 DR HSP; P13163; ISXT.  
 DR InterPro: IPR001961; Strep/Strep-toxin.  
 DR Pfam: PF01123; Strep\_Strep-toxin; 1.  
 DR Pfam: PF02876; Strep\_Strep-toxin; 1.  
 DR PRINTS: PR00279; BACTRLTOXIN.  
 DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.  
 FT NON\_TER  
 SQ SEQUENCE 209 AA; 24086 MW; 1C4343235157DB68 CRC64;

Query Match 72.7%; Score 24; DB 2; Length 209;  
 Best Local Similarity 50.0%; Pred. No. 46;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12  
 DB 125 KTTVTAQAEID 134

RESULT 11  
 Q91AD2 PRELIMINARY; PRT; 209 AA.  
 AC Q91AD2;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Mitogenic exotoxin Z-13 (Fragment).  
 GN SMEZ-13.  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
 OC Streptococcaceae; Streptococcus.  
 OX NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=95/31;  
 RX MEDLINE=20273982; PubMed=10811869;  
 RA Prof. T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,  
 Fraser J.D.;  
 RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,  
 Mosaic Structure, and Significant Antigenic Variation.";  
 RL J. Exp. Med. 191:1765-1776(2000).  
 DR EMBL; AF143663; AAF66663.1; -.  
 DR HSP; P13163; ISXT.  
 DR InterPro: IPR001961; Strep/Strep-toxin.  
 DR Pfam: PF01123; Strep\_Strep-toxin; 1.  
 DR Pfam: PF02876; Strep\_Strep-toxin; 1.  
 DR PRINTS: PR00279; BACTRLTOXIN.  
 DR PROSITE: PS00278; STAPH\_STREP\_TOXIN\_2; 1.  
 FT NON\_TER  
 SQ SEQUENCE 209 AA; 24072 MW; 984C1B4614589A1E CRC64;

Query Match 72.7%; Score 24; DB 2; Length 209;  
 Best Local Similarity 50.0%; Pred. No. 46;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12  
 DB 125 KTTVTAQAEID 134

RESULT 12  
 Q91AD1 PRELIMINARY; PRT; 209 AA.  
 AC Q91AD1;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Mitogenic exotoxin Z-14 (Fragment).  
 GN SMEZ-14.  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
 OC Streptococcaceae; Streptococcus.  
 OX NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=4202;  
 RX MEDLINE=20273982; PubMed=10811869;  
 RA Prof. T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,  
 Fraser J.D.;  
 RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,  
 Mosaic Structure, and Significant Antigenic Variation.";  
 RL J. Exp. Med. 191:1765-1776(2000).  
 DR EMBL; AF143664; AAF66664.1; -.  
 DR HSP; P13163; ISXT.  
 DR InterPro: IPR001961; Strep/Strep-toxin.  
 DR Pfam: PF01123; Strep\_Strep-toxin; 1.  
 DR Pfam: PF02876; Strep\_Strep-toxin; 1.

DR PRINTS; PRO0279; BACTRITOXIN.  
 DR PROSITE; PS00278; STAPH\_STREP\_TOXIN\_2; 1.  
 RT NON\_TER 1  
 SQ SEQUENCE 209 AA; 24122 MW; 0CF5D429E1B96EFE CRC64;

Query Match 72.7%; Score 24; DB 2; Length 209;  
 Best Local Similarity 50.0%; Pred. No. 46;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
 | | | |  
 Db 125 KTTVTAQED 134

RESULT 13  
 O9LAD0 PRELIMINARY; PRT; 209 AA.

AC O9LAD0; 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Mitogenic exotoxin z-15 (Fragment).  
 GN SMEZ-15.

OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
 OC Streptococcaceae; Streptococcus.

OX NCBI\_TaxID=1314;  
 [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-1;  
 RX MEDLINE=20273982; PubMed=10811869;

RA Profit T., Mofatt S.L., Weller K.D., Paterson A., Martin D.,  
 RA Fraser J.D.;

RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,  
 RT Mosaic Structure, and Significant Antigenic Variation.";  
 RL J. Exp. Med. 191:1765-1776(2000).

DR EMBL; AF143665; AAF6665.1; -.

DR HSSP; P13163; ISXT.

DR InterPro; IPR001961; Staph/Strep\_toxin.

DR Pfam; PF01123; Staph\_Strep\_toxin; 1.

DR Pfam; PF02876; Staph\_Strep\_tox\_C; 1.

DR PRINTS; PRO0279; BACTRITOXIN.

DR PROSITE; PS00278; STAPH\_STREP\_TOXIN\_2; 1.

FT NON\_TER 1  
 SQ SEQUENCE 209 AA; 24209 MW; B9097B4ABCD1474 CRC64;

Query Match 72.7%; Score 24; DB 2; Length 209;  
 Best Local Similarity 50.0%; Pred. No. 46;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
 | | | |  
 Db 125 KTTVTAQED 134

RESULT 14  
 O9LAC9 PRELIMINARY; PRT; 209 AA.

AC O9LAC9; 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Mitogenic exotoxin z-16 (Fragment).  
 GN SMEZ-16.

OS Streptococcus pyogenes.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
 OC Streptococcaceae; Streptococcus.

OX NCBI\_TaxID=1314;  
 [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-10649;

RX MEDLINE=20273982; PubMed=10811869;

RA Profit T., Mofatt S.L., Weller K.D., Paterson A., Martin D.,

RA Fraser J.D.;  
 RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,  
 RT Mosaic Structure, and Significant Antigenic Variation.";  
 RL J. Exp. Med. 191:1765-1776(2000).

DR EMBL; AF143666; AAF6666.1; -.

DR HSSP; P13163; ISXT.

DR InterPro; IPR001961; Staph/Strep\_toxin.

DR Pfam; PF01123; Staph\_Strep\_toxin; 1.

DR Pfam; PF02876; Staph\_Strep\_tox\_C; 1.

DR PRINTS; PRO0279; BACTRITOXIN.

DR PROSITE; PS00278; STAPH\_STREP\_TOXIN\_2; 1.

FT NON\_TER 1  
 SQ SEQUENCE 209 AA; 24172 MW; A1DB8FA187098BA5 CRC64;

Query Match 72.7%; Score 24; DB 2; Length 209;  
 Best Local Similarity 50.0%; Pred. No. 46;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
 | | | |  
 Db 125 KTTVTAQED 134

RESULT 15  
 O9LAC8 PRELIMINARY; PRT; 209 AA.

AC O9LAC8; 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Mitogenic exotoxin z-17 (Fragment).  
 GN SMEZ-17.

OS Streptococcus pyogenes.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
 OC Streptococcaceae; Streptococcus.

OX NCBI\_TaxID=1314;  
 [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-11686;

RX MEDLINE=20273982; PubMed=10811869;

RA Profit T., Mofatt S.L., Weller K.D., Paterson A., Martin D.,  
 RA Fraser J.D.;

RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,  
 RT Mosaic Structure, and Significant Antigenic Variation.";  
 RL J. Exp. Med. 191:1765-1776(2000).

DR EMBL; AF143667; AAF6667.1; -.

DR HSSP; P13163; ISXT.

DR InterPro; IPR001961; Staph/Strep\_toxin.

DR Pfam; PF01123; Staph\_Strep\_toxin; 1.

DR Pfam; PF02876; Staph\_Strep\_tox\_C; 1.

DR PRINTS; PRO0279; BACTRITOXIN.

DR PROSITE; PS00278; STAPH\_STREP\_TOXIN\_2; 1.

FT NON\_TER 1  
 SQ SEQUENCE 209 AA; 24181 MW; A33BF466398D9DC2 CRC64;

Query Match 72.7%; Score 24; DB 2; Length 209;  
 Best Local Similarity 50.0%; Pred. No. 46;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
 | | | |  
 Db 125 KTTVTAQED 134

Search completed: June 5, 2003, 15:55:45  
 Job time : 35.9091 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:47:43 ; Search time 40.7273 Seconds  
(without alignments)  
39.261 Million cell updates/sec

```
Title: US-09-150-947F-17
Perfect score: 28
Sequence: 1 XXXXXXXXXXXXLD 12
```

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq.101002.\*

1:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*

**Pred.** No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	19	67.9	378	21	AA635560	Arabidopsis thaliana
2	18	64.3	46	21	AAB38172	Gene 3 human secret
3	18	64.3	46	21	AAB38173	Human secreted pro
4	18	64.3	147	21	AA603207	Human secreted pro
5	18	64.3	152	19	AAW48271	Rat ninjurin 1, R
6	18	64.3	152	19	AAW48272	Human ninjurin 1,
7	18	64.3	173	22	ABG20269	Novel human diagno
8	18	64.3	200	19	AAW74903	Human secreted pro
9	18	64.3	213	22	AAW47003	Plant D-like cyclin
10	18	64.3	247	13	AAK29641	PCD1 ORF 6, Chlam

11	18	64.3	284	23	ABP11256
12	18	64.3	406	17	AAW33144
13	18	64.3	420	17	AAW33155
14	18	64.3	420	17	AAW33157
15	18	64.3	618	22	AAW19070
16	18	64.3	705	17	AAW068311
17	18	64.3	765	22	AAW36337
18	18	64.3	755	22	AAW437373
19	18	64.3	768	21	AAW24448
20	17	60.7	17	23	ABW16254
21	17	60.7	17	23	ABW16259
22	17	60.7	24	21	AAW97855
23	17	60.7	24	21	AAW97852
24	17	60.7	24	22	AAW72183
25	17	60.7	24	22	AAW72200
26	17	60.7	75	23	ABW1309
27	17	60.7	78	22	ABG15641
28	17	60.7	88	22	AAW11653
29	17	60.7	99	21	AAW35153
30	17	60.7	100	21	AAW32653
31	17	60.7	106	23	ABP32548
32	17	60.7	142	19	AAW48273
33	17	60.7	142	20	AAW28249
34	17	60.7	142	21	AAW01389
35	17	60.7	142	21	AAW39036
36	17	60.7	142	23	ABW55755
37	17	60.7	143	22	ABW12218
38	17	60.7	150	23	AAW36961
39	17	60.7	153	21	AAW60286
40	17	60.7	170	23	ABW80300
41	17	60.7	171	21	AAW35162
42	17	60.7	184	22	ABW66316
43	17	60.7	186	21	ABW22925
44	17	60.7	186	21	AAW60286
45	17	60.7	189	21	AAW60295

## ALIGNMENTS

## RESULT 1

ID AAG35560 standard; protein; 378 AA.

AC AAG35560;

DT 18-OCT-2000 (flrst entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 43459

KW Protein identification; signal transduction pathway; metabolic pathway; hub identification

KW termination sequence.

05 *Arabidopsis thaliana*.

PN EP1033405-A2

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 09-MAR-1999; 99US-0123548.

PR 25-MAR-1999; 99US-0126264.

PR 01-APR-1999; 99US-0127462.

PR 08-APR-1999; 99US-0128714.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999;	99US-0130449.
PR 23-APR-1999;	99US-0130510.
PR 23-APR-1999;	99US-0130891.
PR 28-APR-1999;	99US-0131449.
PR 30-APR-1999;	99US-0132048.
PR 04-MAY-1999;	99US-0132407.
PR 05-MAY-1999;	99US-0132485.
PR 06-MAY-1999;	99US-0132486.
PR 07-MAY-1999;	99US-0132863.
PR 11-MAY-1999;	99US-0134256.
PR 14-MAY-1999;	99US-0134218.
PR 14-MAY-1999;	99US-0134219.
PR 14-MAY-1999;	99US-0134221.
PR 14-MAY-1999;	99US-0134370.
PR 18-MAY-1999;	99US-0134768.
PR 19-MAY-1999;	99US-0134941.
PR 20-MAY-1999;	99US-0135124.
PR 21-MAY-1999;	99US-0135353.
PR 24-MAY-1999;	99US-0135629.
PR 25-MAY-1999;	99US-0136021.
PR 27-MAY-1999;	99US-0136392.
PR 28-MAY-1999;	99US-0136782.
PR 01-JUN-1999;	99US-0137222.
PR 03-JUN-1999;	99US-0137528.
PR 04-JUN-1999;	99US-0137502.
PR 07-JUN-1999;	99US-0137724.
PR 08-JUN-1999;	99US-0138094.
PR 10-JUN-1999;	99US-0138540.
PR 10-JUN-1999;	99US-0138847.
PR 14-JUN-1999;	99US-0139111.
PR 16-JUN-1999;	99US-0139452.
PR 16-JUN-1999;	99US-0139453.
PR 17-JUN-1999;	99US-0139492.
PR 18-JUN-1999;	99US-0139454.
PR 18-JUN-1999;	99US-0139455.
PR 18-JUN-1999;	99US-0139456.
PR 18-JUN-1999;	99US-0139457.
PR 18-JUN-1999;	99US-0139458.
PR 18-JUN-1999;	99US-0139459.
PR 18-JUN-1999;	99US-0139460.
PR 18-JUN-1999;	99US-0139461.
PR 18-JUN-1999;	99US-0139462.
PR 18-JUN-1999;	99US-0139463.
PR 18-JUN-1999;	99US-0139750.
PR 21-JUN-1999;	99US-0139763.
PR 22-JUN-1999;	99US-0139817.
PR 23-JUN-1999;	99US-0139899.
PR 23-JUN-1999;	99US-0140353.
PR 23-JUN-1999;	99US-0140354.
PR 24-JUN-1999;	99US-0140695.
PR 28-JUN-1999;	99US-0140823.
PR 29-JUN-1999;	99US-0140991.
PR 30-JUN-1999;	99US-0141287.
PR 01-JUL-1999;	99US-0141842.
PR 01-JUL-1999;	99US-0142153.
PR 02-JUL-1999;	99US-0142053.
PR 06-JUL-1999;	99US-0142390.
PR 08-JUL-1999;	99US-0142803.
PR 09-JUL-1999;	99US-0142920.
PR 12-JUL-1999;	99US-0142977.
PR 13-JUL-1999;	99US-0143544.
PR 14-JUL-1999;	99US-0143624.
PR 15-JUL-1999;	99US-0144005.
PR 16-JUL-1999;	99US-0144086.
PR 16-JUL-1999;	99US-0144086.
PR 19-JUL-1999;	99US-0144325.
PR 19-JUL-1999;	99US-0144331.
PR 19-JUL-1999;	99US-0144332.
PR 19-JUL-1999;	99US-0144333.
PR 19-JUL-1999;	99US-0144334.
PR 19-JUL-1999;	99US-0144335.
PR 20-JUL-1999;	99US-0144352.
PR 20-JUL-1999;	99US-0144632.
PR 21-JUL-1999;	99US-0144884.
PR 21-JUL-1999;	99US-0144814.
PR 21-JUL-1999;	99US-0145086.
PR 21-JUL-1999;	99US-0145088.
PR 22-JUL-1999;	99US-0145085.
PR 22-JUL-1999;	99US-0145087.
PR 22-JUL-1999;	99US-0145089.
PR 22-JUL-1999;	99US-0145192.
PR 23-JUL-1999;	99US-0145145.
PR 23-JUL-1999;	99US-0145218.
PR 23-JUL-1999;	99US-0145224.
PR 26-JUL-1999;	99US-0145276.
PR 27-JUL-1999;	99US-0145913.
PR 27-JUL-1999;	99US-0145918.
PR 27-JUL-1999;	99US-0145919.
PR 28-JUL-1999;	99US-0145951.
PR 02-AUG-1999;	99US-0146386.
PR 02-AUG-1999;	99US-0146388.
PR 02-AUG-1999;	99US-0146389.
PR 03-AUG-1999;	99US-0147038.
PR 04-AUG-1999;	99US-0147204.
PR 05-AUG-1999;	99US-0147302.
PR 05-AUG-1999;	99US-0147192.
PR 05-AUG-1999;	99US-0147260.
PR 06-AUG-1999;	99US-0147303.
PR 06-AUG-1999;	99US-0147416.
PR 09-AUG-1999;	99US-0147493.
PR 09-AUG-1999;	99US-0147935.
PR 10-AUG-1999;	99US-0148171.
PR 11-AUG-1999;	99US-0148319.
PR 12-AUG-1999;	99US-0148341.
PR 13-AUG-1999;	99US-0148565.
PR 16-AUG-1999;	99US-0148684.
PR 17-AUG-1999;	99US-0149368.
PR 18-AUG-1999;	99US-0149175.
PR 20-AUG-1999;	99US-0149426.
PR 20-AUG-1999;	99US-0149722.
PR 20-AUG-1999;	99US-0149723.
PR 20-AUG-1999;	99US-0149929.
PR 23-AUG-1999;	99US-0149902.
PR 23-AUG-1999;	99US-0149930.
PR 23-AUG-1999;	99US-0150566.
PR 26-AUG-1999;	99US-0150884.
PR 27-AUG-1999;	99US-0151065.
PR 27-AUG-1999;	99US-0151066.
PR 27-AUG-1999;	99US-0151080.
PR 30-AUG-1999;	99US-0151303.
PR 31-AUG-1999;	99US-0151348.
PR 01-SEP-1999;	99US-0151930.
PR 07-SEP-1999;	99US-0152363.
PR 10-SEP-1999;	99US-0153070.
PR 13-SEP-1999;	99US-0153758.
PR 15-SEP-1999;	99US-0154018.
PR 16-SEP-1999;	99US-0154039.
PR 20-SEP-1999;	99US-0154779.
PR 22-SEP-1999;	99US-0155139.
PR 23-SEP-1999;	99US-0155486.
PR 24-SEP-1999;	99US-0155659.
PR 28-SEP-1999;	99US-0156458.
PR 29-SEP-1999;	99US-0156596.
PR 04-OCT-1999;	99US-0157117.
PR 05-OCT-1999;	99US-0157753.
PR 06-OCT-1999;	99US-0157865.
PR 07-OCT-1999;	99US-0158029.
PR 08-OCT-1999;	99US-0158232.
PR 12-OCT-1999;	99US-0158369.
PR 13-OCT-1999;	99US-0159293.
PR 13-OCT-1999;	99US-0159294.
PR 13-OCT-1999;	99US-0159325.
PR 14-OCT-1999;	99US-0159329.
PR 14-OCT-1999;	99US-0159330.

PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 67.9%; Score 19; DB 21; Length 378;  
Best Local Similarity 40.0%; Pred. No. 4.9e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KXXXXXXMD 12  
DB 13 KKAADAAALD 22

RESULT 2  
AAB38172  
ID AAB38172 standard; Protein; 46 AA.  
XX  
AC AAB38172;  
XX  
DT 30-JAN-2001 (first entry)  
XX  
DE Gene 3 human secreted protein homologous amino acid sequence #111.  
XX

KW Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;  
KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;  
KW cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;  
KW fungicide; ophthalmological; gene therapy; autoimmune disease; infection;  
KW hyperproliferative disorder; cardiovascular disorder; anglogenesis;  
KW cerebrovascular disorder; nervous system disorder; ocular disorder;  
KW wound healing; skin aging; food additive; preservative.

OS Homo sapiens.  
XX  
PN WO200058468-A2.  
XX  
PD 05-OCT-2000.  
XX  
PF 22-MAR-2000; 2000MO-US07526.  
XX  
PR 26-MAR-1999; 99US-0126600.  
PR 22-DEC-1999; 99US-0171550.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX

PI Rosen CA, Ruben SM, Komatsoulis G;  
XX  
XX WPI; 2000-611713/58.  
XX

PT Nucleic acids encoding human secreted proteins, used to prevent, treat,  
PT ameliorate, or diagnose conditions such as autoimmune disorders, skin  
PT disorders and cancer -  
XX  
XX

PS Disclosure; Page 362; 374pp; English.

XX The polynucleotide sequences given in AAC69399 to AAC69445 encode the  
CC human secreted proteins given in AAB38119 to AAB38165. AAB38166 to  
CC AAB38201 represent human secreted polypeptide sequences and proteins  
CC homologous to them, which are given in the exemplification of the present  
CC invention. Human secreted proteins have activities based on the tissues  
CC and cells the genes are expressed in. Example of activities include:  
CC immunosuppressive; antiarthritic; antirheumatic; antiproliferative;  
CC cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;  
CC neuroprotective; antibacterial; virucide; fungicide; and  
CC ophthalmological. The polynucleotides and polypeptides can be used to  
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
CC in diagnosing a pathological condition or susceptibility to a  
CC pathological condition. Disorders which are diagnosed or treated include  
CC autoimmune diseases, hyperproliferative disorders, cardiovascular  
CC disorders, cerebrovascular disorders, anglogenesis, nervous system  
CC disorders, infections caused by bacteria, viruses and fungi and ocular  
CC disorders. The polypeptides can also be used to aid wound healing and  
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to  
CC maintain organs before transplantation, for supporting cell culture of  
CC primary tissues, to regenerate tissues and in chemotaxis. The  
CC polypeptides can also be used as a food additive or preservative to  
CC increase or decrease storage capabilities. AAC69390 to AAC69398 and  
CC AAB38118 represent sequences used in the exemplification of the present  
CC invention.  
XX

SO Sequence 46 AA;  
Query Match 64.3%; Score 18; DB 21; Length 46;  
Best Local Similarity 40.0%; Pred. No. 1.5e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KXXXXXXMD 12  
DB 19 KKSAAESMDL 28

RESULT 3  
AAB38173  
ID AAB38173 standard; Protein; 46 AA.  
XX  
AC AAB38173;  
XX  
DT 30-JAN-2001 (first entry)  
XX  
DE Human secreted protein sequence encoded by gene 3 SEQ ID NO:112.  
XX

KW Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;  
KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;  
KW cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;  
KW fungicide; ophthalmological; gene therapy; autoimmune disease; infection;  
KW hyperproliferative disorder; cardiovascular disorder; anglogenesis;  
KW cerebrovascular disorder; nervous system disorder; ocular disorder;  
KW wound healing; skin aging; food additive; preservative.

OS Homo sapiens.  
XX  
PN WO200058468-A2.  
XX  
PD 05-OCT-2000.  
XX  
PF 22-MAR-2000; 2000MO-US07526.  
XX  
PR 26-MAR-1999; 99US-0126600.  
PR 22-DEC-1999; 99US-0171550.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX

PI Rosen CA, Ruben SM, Komatsoulis G;  
XX  
XX WPI; 2000-611713/58.  
XX  
XX

PT Nucleic acids encoding human secreted proteins, used to prevent, treat,  
 PT ameliorate, or diagnose conditions such as autoimmune disorders, skin  
 PT disorders and cancer -  
 PS  
 XX Disclosure, Page 362, 374pp; English.

CC The polynucleotide sequences given in AAC69399 to AAC69445 encode the  
 CC human secreted proteins given in AAB38119 to AAB38165. AAB38166 to  
 CC AAB38201 represent human secreted polypeptide sequences and proteins  
 CC homologous to them, which are given in the exemplification of the present  
 CC invention. Human secreted proteins have activities based on the tissues  
 CC and cells the genes are expressed in. Example of activities include:  
 CC immunosuppressive; antiarthritic; antineumatic; antiproliferative;  
 CC cytoprotective; cardiac; vasotropic; cerebroprotective; neurotropic;  
 CC ophthalmological. The polynucleotides and polypeptides can be used to  
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
 CC in diagnosing a pathological condition or susceptibility to a  
 CC pathological condition. Disorders which are diagnosed or treated include  
 CC autoimmune diseases, hyperproliferative disorders, cardiovascular  
 CC disorders, cerebrovascular disorders, angiogenesis, nervous system  
 CC disorders. The polypeptides can also be used to aid wound healing and  
 CC epithelial cell proliferation, to prevent skin aging due to sunburn, to  
 CC maintain organs before transplantation, for supporting cell culture of  
 CC primary tissues, to regenerate tissues and in chemotaxis. The  
 CC polypeptides can also be used as a food additive or preservative and  
 CC increase or decrease storage capabilities. AAC69399 to AAC69398 and  
 CC AAB38118 represent sequences used in the exemplification of the present  
 CC invention.

SQ Sequence 46 AA;

Query Match 64.3%; Score 18; DB 21; Length 46;

Best Local Similarity 40.0%; Pred. No. 1.5e+02;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KKKXXXXXLD 12

DB 19 KKSAAESMMD 28

RESULT 4

AAG02207 ID AAG02207 standard; Protein; 147 AA.

AC AAG02207;

DT 06-OCT-2000 (first entry)

DE Human secreted protein, SEQ ID NO: 6288.

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

KW gene therapy; chromosome mapping.

OS Homo sapiens.

PN EPI033401-A2.

PD 06-SEP-2000.

PF 21-FEB-2000; 2000EP-0200610.

PR 26-FEB-1999; 99US-0122487.

PA (GBST ) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

DR N-PSDB; AAC02213.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 PS  
 XX Claim 13; SEQ ID 6288; 71pp + CD-ROM; English.

CC The present sequence is a polypeptide encoded by one of a large number  
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
 CC were prepared from total human RNAs or poly(A) RNAs derived from 30  
 CC different tissues. EST sequences usually correspond mainly to the 3'  
 CC untranslated region (UTR) of the mRNA because they are often obtained  
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
 CC those cases where longer cDNA sequences have been obtained, the full 5'  
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
 CC ends and can therefore be used to obtain full length cDNAs and genomic  
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
 CC chromosome mapping procedures. They are used to obtain upstream  
 CC regulatory sequences and to design expression and secretion vectors.

SQ Sequence 147 AA;

Query Match 64.3%; Score 18; DB 21; Length 147;

Best Local Similarity 40.0%; Pred. No. 4.4e+02;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KKKXXXXXLD 12

DB 78 KKKRAATITLD 87

RESULT 5

AAW48271 ID AAW48271 standard; Protein; 152 AA.

AC AAW48271;

DT 23-JUN-1998 (first entry)

DE Rat nInjurin 1.

KW Rat; nInjurin; cellular adhesion molecule; membrane bound; tumour;

KW nerve injury induced gene; inflammation; nervous system.

OS Rattus sp.

PN MO9803650-A1.

PD 29-JAN-1998.

PF 24-JUL-1997; 97WO-US12210.

PR 24-JUL-1996; 96US-0672850.

PA (UNITW ) UNITV WASHINGTON.

PI Araki T, MIdrandt J;

DR WPI; 1998-120775/11.

DR N-PSDB; AAV20665.

PT New isolated nerve injury induced (nInjurin) gene - used to develop

PT products for treating conditions involving excessive or insufficient

PT cellular adhesion, e.g. inflammation or tumours

PS Claim 5; Fig 1B; 80pp; English.

CC The present sequence represents rat nInjurin 1 (nerve injury induced).  
 CC The nInjurin protein (NP) plays a role in axonal regeneration of  
 CC peripheral nervous system (PNS) neuronal cells after injury. The  
 CC products can be used for developing products for treating nInjurin  
 CC mediated disorders including conditions involving inappropriate (i.e.  
 CC excessive or insufficient) cellular adhesion. Conditions involving

CC excessive cellular adhesion which may be treated include e.g.  
 CC inflammatory diseases such as rheumatoid arthritis, asthma, allergy  
 CC conditions, adult respiratory distress syndrome, inflammatory bowel  
 CC diseases (e.g. Crohn's disease, ulcerative colitis and regional  
 CC enteritis) and ophthalmic inflammatory diseases, autoimmune diseases,  
 CC thrombosis or inappropriate platelet aggregation conditions,  
 CC arteriosclerosis, reocclusion following thrombolysis, cardiovascular  
 CC metastasis conditions. The products can also be used to promote  
 CC cellular adhesion, e.g. in nerve regeneration, wound healing or  
 CC prosthetic implantation. The products can also be used for detection,  
 CC purification, diagnosis and screening assays.

XX  
 SQ Sequence 152 AA;

QY 3 KKKXXXXXXLD 12  
 II  
 II  
 DB 44 KKSAAESMLD 53

Query Match 64.3%; Score 18; DB 19; Length 152;  
 Best Local Similarity 40.0%; Pred. No. 4.5e+02;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

RESULT 6  
 AAW48272 standard; Protein: 152 AA.

XX AAW48272:  
 AC AAW48272:  
 XX 23-JUN-1998 (first entry)  
 DT  
 XX  
 DE Human nInjurin 1.  
 XX  
 KM Human: nInjurin; cellular adhesion molecule; membrane bound; tumour;  
 KW nerve injury induced gene; inflammation; nervous system.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WC9803650-A1.  
 XX  
 PD 29-JAN-1998.  
 XX  
 PF 24-JUL-1997; 97WO-US12210.  
 XX  
 PR 24-JUL-1996; 96US-0672850.  
 XX  
 PA (UNIV ) UNIV WASHINGTON.  
 XX  
 PI Araki T, Milbrandt J;  
 XX  
 DR WPI: 1998-120775/11.  
 DR N-PSDB; AAV20666.  
 XX  
 PT New isolated nerve injury induced (nInjurin) gene - used to develop  
 PT products for treating conditions involving excessive or insufficient  
 PT cellular adhesion, e.g. inflammation or tumours  
 XX  
 PS Claim 6; Fig 2B; 80pp; English.

CC The present sequence represents human nInjurin 1 (nerve injury induced).  
 CC The nInjurin protein (NP) plays a role in axonal regeneration of  
 CC peripheral nervous system (PNS) neuronal cells after injury. The  
 CC products can be used for developing products for treating nInjurin  
 CC mediated disorders including conditions involving inappropriate (i.e.  
 CC excessive or insufficient) cellular adhesion. Conditions involving  
 CC inflammatory diseases such as rheumatoid arthritis, asthma, allergy  
 CC conditions, adult respiratory distress syndrome, inflammatory bowel  
 CC diseases (e.g. Crohn's disease, ulcerative colitis and regional  
 CC enteritis) or inappropriate platelet aggregation conditions,  
 CC arteriosclerosis, reocclusion following thrombolysis, cardiovascular

CC diseases, some forms of diabetes and neoplastic disease including  
 CC metastasis conditions. The products can also be used to promote  
 CC cellular adhesion, e.g. in nerve regeneration, wound healing or  
 CC prosthetic implantation. The products can also be used for detection,  
 CC purification, diagnosis and screening assays.

XX  
 SQ Sequence 152 AA;

QY 3 KKKXXXXXXLD 12  
 II  
 II  
 DB 44 KKSAAESMLD 53

Query Match 64.3%; Score 18; DB 19; Length 152;  
 Best Local Similarity 40.0%; Pred. No. 4.5e+02;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

RESULT 7  
 ABG20269 standard; Protein: 173 AA.

XX ABG20269:  
 AC ABG20269:  
 XX 18-FEB-2002 (first entry)  
 DT  
 XX  
 DE Novel human diagnostic protein #20260.  
 XX  
 KM Human: chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US08631.  
 XX  
 PR 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Drmanac RT, Liu C, Tang YT;  
 XX  
 DR WPI: 2001-639362/73.  
 DR N-PSDB; AAS84456.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity  
 XX  
 PS Claim 20; SEQ ID No 50628; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIP0 at ftp.wipo.int/pub/published\_pct\_sequences.

CC Sequence 173 AA:

Query Match 64.3%; Score 18; DB 22; Length 173;  
Best Local Similarity 40.0%; Pred. No. 5.1e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KXXXXXXXXXD 12  
11 11  
DB 141 KKAASHRALD 150

# RESULT 8

AAW74903  
ID AAW74903 standard; Protein; 200 AA.

AC AAW74903;

DT 25-JAN-1999 (first entry)

DE Human secreted protein encoded by gene 176 clone HFTBR48.

KW Human; secreted protein; testis; tumour; foetal brain tissue;  
fusion protein; cancer; central nervous system; seizure;  
diagnosis; neurodegenerative disease.

KX Homo sapiens.

OS Homo sapiens.

FT Key Location/Qualifiers

PD Misc-difference 200

PN W09839448-A2.

PF 11-SEP-1998.

PR 06-MAR-1998; 98WO-US04493.

PR 02-OCT-1997; 97US-0061060.

PR 07-MAR-1997; 97US-0038621.

PR 07-MAR-1997; 97US-0040161.

PR 07-MAR-1997; 97US-0040162.

PR 07-MAR-1997; 97US-0040163.

PR 07-MAR-1997; 97US-0040333.

PR 07-MAR-1997; 97US-0040334.

PR 07-MAR-1997; 97US-0040336.

PR 07-MAR-1997; 97US-0040626.

PR 11-APR-1997; 97US-0043311.

PR 11-APR-1997; 97US-0043312.

PR 11-APR-1997; 97US-0043313.

PR 11-APR-1997; 97US-0043314.

PR 11-APR-1997; 97US-0043568.

PR 11-APR-1997; 97US-0043569.

PR 11-APR-1997; 97US-0043576.

PR 11-APR-1997; 97US-0043578.

PR 11-APR-1997; 97US-0043580.

PR 23-MAY-1997; 97US-0047585.

PR 23-MAY-1997; 97US-0047586.

PR 23-MAY-1997; 97US-0047587.

PR 23-MAY-1997; 97US-0047588.

PR 23-MAY-1997; 97US-0047589.

PR 23-MAY-1997; 97US-0047590.

PR 23-MAY-1997; 97US-0047592.

PR 23-MAY-1997; 97US-0047593.

PR 23-MAY-1997; 97US-0047594.

PR 23-MAY-1997; 97US-0047595.

PR 23-MAY-1997; 97US-0047596.

PR 23-MAY-1997; 97US-0047597.

PR 23-MAY-1997; 97US-0047598.

PR 23-MAY-1997; 97US-0047599.

PR 23-MAY-1997; 97US-0047600.

PR 23-MAY-1997; 97US-0047601.

PR 23-MAY-1997; 97US-0047612.

PR 23-MAY-1997; 97US-0047613.

PR 23-MAY-1997; 97US-0047614.

PR 23-MAY-1997; 97US-0047615.

PR 23-MAY-1997; 97US-0047617.

PR 23-MAY-1997; 97US-0047618.

PR 23-MAY-1997; 97US-0047632.

PR 23-MAY-1997; 97US-0047633.

PR 06-JUN-1997; 97US-0048964.

PR 06-JUN-1997; 97US-0048974.

PR 13-JUN-1997; 97US-0049610.

PR 08-JUL-1997; 97US-0051926.

PR 16-JUL-1997; 97US-0052874.

PR 18-AUG-1997; 97US-0055724.

PR 22-AUG-1997; 97US-0056630.

PR 22-AUG-1997; 97US-0056631.

PR 22-AUG-1997; 97US-0056632.

PR 22-AUG-1997; 97US-0056636.

PR 22-AUG-1997; 97US-0056637.

PR 22-AUG-1997; 97US-0056662.

PR 22-AUG-1997; 97US-0056664.

PR 22-AUG-1997; 97US-0056682.

PR 22-AUG-1997; 97US-0056684.

PR 22-AUG-1997; 97US-0056686.

PR 22-AUG-1997; 97US-0056687.

PR 22-AUG-1997; 97US-0056688.

PR 22-AUG-1997; 97US-0056689.

PR 22-AUG-1997; 97US-0056692.

PR 22-AUG-1997; 97US-0056693.

PR 22-AUG-1997; 97US-0056694.

PR 22-AUG-1997; 97US-0056699.

PR 22-AUG-1997; 97US-0056711.

PR 22-AUG-1997; 97US-0056712.

PR 22-AUG-1997; 97US-0056713.

PR 22-AUG-1997; 97US-0056714.

PR 22-AUG-1997; 97US-0056715.

PR 22-AUG-1997; 97US-0056716.

PR 22-AUG-1997; 97US-0056717.

PR 22-AUG-1997; 97US-0056718.

PR 22-AUG-1997; 97US-0056719.

PR 22-AUG-1997; 97US-0056880.

PR 22-AUG-1997; 97US-0056881.

PR 22-AUG-1997; 97US-0056882.

PR 22-AUG-1997; 97US-0056884.

PR 22-AUG-1997; 97US-0056886.

PR 22-AUG-1997; 97US-0056887.

PR 22-AUG-1997; 97US-0056888.

PR 22-AUG-1997; 97US-0056889.

PR 22-AUG-1997; 97US-0056892.

PR 22-AUG-1997; 97US-0056893.

PR 22-AUG-1997; 97US-0056894.

PR 22-AUG-1997; 97US-0056910.

PR 22-AUG-1997; 97US-0056911.

PR 05-SEP-1997; 97US-0057650.

PR 05-SEP-1997; 97US-0057659.

PR 05-SEP-1997; 97US-0057761.

PR 12-SEP-1997; 97US-0058785.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA;

PI Feng P, Ferrie AM, Fischer CL, Florence KA, Greene JM, Hu JS;

PI Kyaw H, Latleur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA;

PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;  
 XX WPI: 1998-506364/43.  
 DR N-PSDB; AAV59686.  
 XX  
 PT New isolated human genes and the secreted polypeptide(s) they encode  
 PT - useful for diagnosis and treatment of e.g. cancers, neurological  
 PT disorders, immune diseases, inflammation or blood disorders  
 XX  
 PS Claim 1; Page 649-650; 721pp; English.  
 XX  
 CC This sequence represents a secreted human protein encoded by the nucleic  
 CC acid molecule designated Gene 176 from the human cDNA clone HPTBR48  
 CC (deposited as clone ATCC 97904 and ATCC 209050).  
 CC The gene can be used to generate fusion proteins by linking to the gene  
 CC to a human immunoglobulin Fc portion (e.g. AAV59502) for increasing the  
 CC stability of the fused protein as compared to the human protein only.  
 CC The invention relates to 186 novel genes and their fragments (nucleic  
 CC acid sequences: AAV59511-V59812; amino acid sequences AAW4731-W5026)  
 CC which are useful for preventing, treating or ameliorating medical  
 CC conditions e.g. by protein or gene therapy. Also, pathological  
 CC conditions can be diagnosed by determining the amount of the new  
 CC polypeptides in a sample or by determining the presence of mutations in  
 CC the new polynucleotides. Specific uses are described for each of the 186  
 CC polynucleotides, based on which tissues they are most highly expressed in  
 CC (see AAV59511 for described uses).  
 CC  
 SQ Sequence 200 AA;  
 XX  
 XX  
 Query Match 64.3%; Score 18; DB 19; Length 200;  
 Best Local Similarity 40.0%; Pred. No. 5.8e+02;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 OY 3 KKKXXXXXLD 12  
 || ||  
 DB 151 KKAASHRAID 160  
 XX  
 XX  
 RESULT 9  
 AAB47003  
 ID AAB47003 standard; Protein; 213 AA.  
 AC AAB47003;  
 XX  
 XX 22-MAR-2001 (first entry)  
 DT  
 XX  
 DE Plant D-like cyclin inhibitor BRO2.  
 XX  
 KW Plant; D-like cyclin inhibitor gene; BRO4; hyperplastic; variant;  
 KW growth rate; dividing cells; inactivation; protoplast; seed;  
 KW root cell; meristem; leaf.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT Protein 20..147  
 FT /Label= BRO2  
 XX  
 XX  
 PN MO200069883-A1.  
 XX  
 PD 23-NOV-2000.  
 XX  
 PF 15-MAY-2000; 2000MO-US13379.  
 XX  
 PR 14-MAY-1999; 99US-0134373.  
 XX  
 PA (HUTC-) HUTCHINSON CANCER RES CENT FRED.  
 XX  
 PI Roberts J, Kelly B;  
 XX  
 XX WPI: 2001-024998/03.  
 DR N-PSDB; AAC85203.  
 XX

PT Functionally inactivating expression of plant D-like cyclin inhibitor  
 PT gene for producing a hyperplastic variant plant, modulating the growth  
 PT and/or yield of plants, and increasing the proportion of dividing cells  
 XX  
 PS Example 1; Page 38-39; 50pp; English.  
 XX  
 CC The sequence given in AAB47003 represents a plant D1 cyclin  
 CC inhibitor, BRO2. The DNA encoding this sequence was isolated using  
 CC a yeast two hybrid screen. The BRO2 protein was found to contain a  
 CC seven amino acid sequence cyclin binding domain similar to that of  
 CC BRO1, BRO3 and BRO4 (See also AAB47005-6). The DNA encoding this  
 CC sequence is homologous to a sequence present in a D-like cyclin  
 CC inhibitor gene and when integrated at the corresponding locus,  
 CC functionally inactivates plant D-like cyclin inhibitor protein  
 CC expression. The BRO4 coding sequence may be used to produce a  
 CC hyperplastic variant plant, increase the growth rate of a plant,  
 CC or increase the proportion of dividing cells in a plant cell  
 CC population, relative to a wild-type plant, by functionally  
 CC inactivating the expression of a plant D-like cyclin inhibitor  
 CC gene in a plant. BRO4 is useful for increasing the proportion of  
 CC dividing cells in a plant cell population comprising protoplast,  
 CC seeds, root cells, meristem cells or leaf cells.  
 CC  
 SQ Sequence 213 AA;  
 XX  
 XX  
 Query Match 64.3%; Score 18; DB 22; Length 213;  
 Best Local Similarity 40.0%; Pred. No. 6.1e+02;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 OY 3 KKKXXXXXLD 12  
 || ||  
 DB 10 KKKRYTSLD 19  
 XX  
 XX  
 RESULT 10  
 AAR29641  
 ID AAR29641 standard; Protein; 247 AA.  
 AC AAR29641;  
 XX  
 XX 05-FEB-1993 (first entry)  
 DT  
 XX  
 DE PCTD ORF 6.  
 XX  
 KW CT; PCTD; epithelium; ocula mucosa; uro-genital mucosa; antigen;  
 KW monoclonal; polyclonal; antibody; vaccine.  
 XX  
 OS Chlamydia trachomatis.  
 XX  
 XX  
 PN EP499681-A.  
 XX  
 PD 26-AUG-1992.  
 XX  
 PF 17-APR-1991; 91EP-0106110.  
 XX  
 PR 07-FEB-1991; 91IT-0000314.  
 XX  
 XX  
 PA (ISTS ) SCLAVO SPA.  
 XX  
 XX Comanducci M, Giuliani MM, Ratti G, Tecce MF;  
 PI WPI: 1992-285922/35.  
 DR N-PSDB; AAQ27429.  
 XX  
 PT PCTD plasmid from Chlamydia Trachomatis and immunogenic proteins  
 PT - for diagnosing and vaccinating against Chlamydia infections  
 PT e.g. venereal lymphogranuloma  
 XX  
 PS Claim 1; Page 8-16; 40pp; English.  
 XX  
 CC The sequences given in AAR29636-43 are encoded by the plasmid isolated  
 CC from Chlamydia trachomatis (CT) serotype D, PCTD. This serotype  
 CC generally infects epithelial tissues, such as the ocular and

uro-genital mucous membranes, and shows a low virulence. Of the eight proteins encoded by the plasmid, seven are encoded by the sense strand and the eighth is encoded by the complementary strand. These proteins can be used as antigens for the preparation of poly- and mono-clonal antibodies to be used in diagnostics. The antigens can also be used in the formulation of vaccines against infections due to CT.

Sequence 247 AA:

Query Match 64.3%; Score 18; DB 13; Length 247;  
Best Local Similarity 40.0%; Pred. No. 7e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KKKXXXXXLD 12  
||  
||  
DB 13 KKNQTAASLD 22

RESULT 11

ABP41256  
ID ABP41256 standard; Protein; 284 AA.

AC ABP41256;

DT 22-AUG-2002 (first entry)

DE Human ovarian antigen HOFNY16, SEQ ID NO:2388.

Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
ovarian cancer; breast cancer; tumour; reproductive system disorder;  
infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;  
inflammatory condition; immune disorder; blood disorder;  
cardiovascular disorder; respiratory disorder; neurological disorder;  
gastrointestinal disorder; urinary system disorder; drug screening;  
gene therapy; chromosome mapping; forensic analysis;  
antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
antiinflammatory; gynaecological; reproductive.

OS Homo sapiens.

PN WO200200677-A1.

PD 03-JAN-2002.

PE 07-JUN-2001; 2001WO-US18569.

PR 07-JUN-2000; 2000US-209467P.

PA (HOMA-) HUMAN GENOME SCI INC.

PI Birse CE, Rosen CA;

DR WPI: 2002-147878/19.

DR N-PSDB; ABO54333.

Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
useful in the prevention, treatment and diagnosis of cancer (e.g.  
neurological diseases -

Claim 11; SEQ ID NO 2388; 2922pp; English.

The invention relates to 2175 novel human ovarian antigens (ABP41054-  
ABP43228) and to cDNAs encoding them (ABO54131-ABO56305), and also  
encompasses polypeptides 90% identical and polynucleotides 95% identical  
to the sequences of the invention. The invention additionally relates to  
recombinant vectors and host cells comprising human ovarian antigen  
polynucleotides, antibodies against human ovarian antigens, and the use  
of ovarian antigen polynucleotides and polypeptides in diagnosing,  
treating, prognosing or preventing various ovarian and/or breast-related  
disorders. Such conditions include ovarian cancer and breast cancer, and  
metastatic tumours of ovarian or breast origin, reproductive system

disorders (e.g., infertility, disorders of pregnancy, anovulation,  
polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine  
disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
vaginitis), immune disorders (e.g., congenital and acquired  
immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
blood-related disorders (e.g., anaemia), cardiovascular disorders,  
respiratory disorders, neurological disorders, gastrointestinal disorders  
and urinary system disorders. Ovarian antigen polypeptides and  
polynucleotides may also be used in screening for compounds which  
modulate ovarian antigen expression or activity. The polynucleotides may  
further be used for gene therapy, chromosome mapping, in the  
identification of individuals and in forensic analysis, and the  
polypeptides may be used as food additives or to prepare antibodies  
useful in disease diagnosis, drug targeting and phenotyping. The present  
sequence represents a human ovarian antigen of the invention. The present  
Note: The sequence data for this patent did not form part of the printed  
specification, but was obtained in electronic format directly from WIPO  
at ftp.wipo.int/pub/published\_pcl\_sequences.

Sequence 284 AA:

Query Match 64.3%; Score 18; DB 23; Length 284;  
Best Local Similarity 40.0%; Pred. No. 7.9e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KKKXXXXXLD 12  
||  
||  
DB 236 KKAASHRALD 245

RESULT 12

AAM93144  
ID AAM93144 standard; Protein; 406 AA.

AC AAM93144;

DT 21-MAY-1999 (first entry)

DE S. lividans xylanase A protein StmXlnA.

Xylanase A; acidophilic; thermostable; XYL I; XYL II; plant biomass;  
hemicellulase; beta-1,4 bond; xylosic chain; xylan; D-xylose; paper;  
pulp; chlorine bleaching; feed; beta-glucan; cellulose; lignin.

OS Streptomyces lividans.

PN US5871730-A.

PD 16-FEB-1999.

PE 29-JUL-1994; 94US-0282197.

PR 29-JUL-1994; 94US-0282197.

PA (UTSH ) UNIV SHERBROOKE.

PI Beaulieu C, Brzezinski R, Dery CV;

DR WPI: 1996-141348/15.

New acidophilic and thermostable xylanase enzymes from Actinomadura  
sp. FC7 - useful for treating plant biomass, especially paper and  
wood pulp, to degrade hemicellulose and hydrolyse xylan

Example 8; Fig 10; 60pp; English.

This invention describes the use of novel acidophilic and thermostable  
xylanase enzymes (XYL I and XYL II) from Actinomadura sp. FC7 which  
retain their activity under harsh industrial conditions (e.g. high  
temperature or wide pH ranges) and may be secreted by recombinant host  
cells, to treat plant biomass. Xylanases XYL I and XYL II are part of  
a large group of hemicellulase enzymes and function by cutting the



CC beta-1,4 bonds within the xylosic chain of xylan (a polymer of D-xylose  
CC residues that is a major constituent of hemicellulose). This means that  
CC they may be used in the paper and pulp industry to improve the efficiency  
CC of the bleaching process by degrading the structure of the material.  
CC XYL I and XYL II may also be used to treat feed, by degrading a  
CC substrate with a high beta-glucan or cellulose content. XYL I and XYL II  
CC retain their activity at high temperatures (e.g. 70 deg. C) and at low  
CC pHs (e.g. 4.0), conditions which tend to denature most known xylanases.  
CC Enzymes that remain active in these conditions may be used in industrial  
CC processes that are carried out at high temperature and low pH to speed up  
CC other, non-enzymatic reactions, minimising costs, energy requirements,  
CC and the risk of pollution, (e.g. enzymes XYL I and XYL II can be used to  
CC facilitate chlorine bleaching of paper pulp which is carried out in hot,  
CC acidic conditions). Pretreatment with XYL I and XYL II, allows the  
CC bleaching agents to penetrate better, to remove lignin from the pulp and  
CC 'bleach' the colouration from it. This means smaller quantities of the  
CC agents can be used to produce the same or a better result. Also,  
CC disrupting the structure aids water drainage.  
CC NOTE: This patent is an equivalent to FI9503640.

SQ Sequence 406 AA;

Query Match 64.3%; Score 18; DB 17; Length 406;  
Best Local Similarity 40.0%; Pred. No. 1.1e+03;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KKKXXXXXLD 12  
|| ||  
Db 330 KKAAYTAVLD 339

RESULT 13  
AAW93155

ID AAW93155 standard; Protein; 420 AA.

XX AAW93155;

DT 21-MAY-1999 (first entry)

XX Streptomyces lividans xylanase A protein.

KW Xylanase A; acidophilic; thermostable; XYL I; XYL II; plant biomass;  
KW hemicellulase; beta-1,4 bond; xylosic chain; xylan; D-xylose; paper;  
KW pulp; chlorine bleaching; feed; beta-glucan; cellulose; lignin.

XX Streptomyces lividans.

PN US5871730-A.

PD 16-FEB-1999.

PF 29-JUL-1994; 94US-0282197.

PR 29-JUL-1994; 94US-0282197.

XX (UYSH ) UNIV SHERBROOKE.

PI Beaulieu C, Brzezinski R, Dery CV;

DR WPI; 1996-141348/15.

PT New acidophilic and thermostable xylanase enzymes from Actinomyadura  
PT sp. FC7 - useful for treating plant biomass, especially paper and  
PT wood pulp, to degrade hemicellulose and hydrolyse xylan

PS Example 8; Fig 11; 60pp; English.

CC This invention describes the use of novel acidophilic and thermostable  
CC xylanase enzymes (XYL I and XYL II) from Actinomyadura sp. FC7 which  
CC retain their activity under harsh industrial conditions (e.g. high  
CC temperature or wide pH ranges) and may be secreted by recombinant host  
CC cells, to treat plant biomass. Xylanases XYL I and XYL II are part of  
CC a large group of hemicellulase enzymes and function by cutting the

CC beta-1,4 bonds within the xylosic chain of xylan (a polymer of D-xylose  
CC residues that is a major constituent of hemicellulose). This means that  
CC they may be used in the paper and pulp industry to improve the efficiency  
CC of the bleaching process by degrading the structure of the material.  
CC XYL I and XYL II may also be used to treat feed, by degrading a  
CC substrate with a high beta-glucan or cellulose content. XYL I and XYL II  
CC retain their activity at high temperatures (e.g. 70 deg. C) and at low  
CC pHs (e.g. 4.0), conditions which tend to denature most known xylanases.  
CC Enzymes that remain active in these conditions may be used in industrial  
CC processes that are carried out at high temperature and low pH to speed up  
CC other, non-enzymatic reactions, minimising costs, energy requirements,  
CC and the risk of pollution, (e.g. enzymes XYL I and XYL II can be used to  
CC facilitate chlorine bleaching of paper pulp which is carried out in hot,  
CC acidic conditions). Pretreatment with XYL I and XYL II, allows the  
CC bleaching agents to penetrate better, to remove lignin from the pulp and  
CC 'bleach' the colouration from it. This means smaller quantities of the  
CC agents can be used to produce the same or a better result. Also,  
CC disrupting the structure aids water drainage.  
CC NOTE: This patent is an equivalent to FI9503640.

SQ Sequence 420 AA;

Query Match 64.3%; Score 18; DB 17; Length 420;  
Best Local Similarity 40.0%; Pred. No. 1.1e+03;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KKKXXXXXLD 12  
|| ||  
Db 330 KKAAYTAVLD 339

RESULT 14  
AAW93157

ID AAW93157 standard; Protein; 420 AA.

XX AAW93157;

DT 21-MAY-1999 (first entry)

XX Streptomyces lividans xylanase A protein.

KW Xylanase A; acidophilic; thermostable; XYL I; XYL II; plant biomass;  
KW hemicellulase; beta-1,4 bond; xylosic chain; xylan; D-xylose; paper;  
KW pulp; chlorine bleaching; feed; beta-glucan; cellulose; lignin.

XX Streptomyces lividans.

PN US5871730-A.

PD 16-FEB-1999.

PF 29-JUL-1994; 94US-0282197.

PR 29-JUL-1994; 94US-0282197.

XX (UYSH ) UNIV SHERBROOKE.

PI Beaulieu C, Brzezinski R, Dery CV;

DR WPI; 1996-141348/15.

PT New acidophilic and thermostable xylanase enzymes from Actinomyadura  
PT sp. FC7 - useful for treating plant biomass, especially paper and  
PT wood pulp, to degrade hemicellulose and hydrolyse xylan

PS Example 9; Fig 13; 60pp; English.

CC This invention describes the use of novel acidophilic and thermostable  
CC xylanase enzymes (XYL I and XYL II) from Actinomyadura sp. FC7 which  
CC retain their activity under harsh industrial conditions (e.g. high  
CC temperature or wide pH ranges) and may be secreted by recombinant host  
CC cells, to treat plant biomass. Xylanases XYL I and XYL II are part of  
CC a large group of hemicellulase enzymes and function by cutting the

CC beta-1,4 bonds within the xylosic chain of xylan (a polymer of D-xylose  
 CC residues that is a major constituent of hemicellulose). This means that  
 CC they may be used in the paper and pulp industry to improve the efficiency  
 CC of the bleaching process by degrading the structure of the material.  
 CC XYL I and XYL II may also be used to treat feed, by degrading a  
 CC substrate with a high beta-glucan or cellulose content. XYL I and XYL II  
 CC retain their activity at high temperatures (e.g. 70 deg. C) and at low  
 CC pHs (e.g. 4.0), conditions which tend to denature most known xylanases.  
 CC Enzymes that remain active in these conditions may be used in industrial  
 CC processes that are carried out at high temperature and low pH to speed up  
 CC other, non-enzymatic reactions, minimising costs, energy requirements,  
 CC and the risk of pollution, (e.g. enzymes XYL I and XYL II can be used to  
 CC facilitate chlorine bleaching of paper pulp which is carried out in hot,  
 CC acidic conditions). Pretreatment with XYL I and XYL II, allows the  
 CC bleaching agents to penetrate better, to remove lignin from the pulp and  
 CC 'bleach' the colouration from it. This means smaller quantities of the  
 CC agents can be used to produce the same or a better result. Also,  
 CC disrupting the structure aids water drainage.  
 CC NOTE: This patent is an equivalent to F19503640.

XX SQ Sequence 420 AA;

Query Match Best Local Similarity 64.3%; Score 18; DB 17; Length 420;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KKKXXXXXXLD 12  
 || ||  
 DB 330 KKAATVAVLD 339

RESULT 15

ABG19070 ID ABG19070 standard; Protein; 618 AA.

AC ABG19070;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #19061.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

XX N-PSDB; AAS83257.

XX New isolated polynucleotide and encoded polypeptides, useful in

XX PT diagnostics, forensics, gene mapping, identification of mutations

XX PT responsible for genetic disorders or other traits and to assess

XX PT biodiversity

XX Claim 20; SEQ ID No 49429; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG0377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 618 AA;

Query Match Best Local Similarity 64.3%; Score 18; DB 22; Length 618;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KKKXXXXXXLD 12  
 || ||  
 DB 372 KKVTLSTSLD 381

Search completed: June 5, 2003, 16:00:25

Job time : 40.7273 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:23:33 ; Search time 10.5455 Seconds  
(without alignments)  
33.481 Million cell updates/sec

Title: US-09-150-947f-17  
Perfect score: 28  
Sequence: 1 XXXXXXXXXXXXLD 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents\_AA:\*  
1: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/5B\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/PCBUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	18	64.3	100	4	US-08-672-850-13
2	18	64.3	152	4	US-08-672-850-4
3	18	64.3	152	4	US-08-672-850-6
4	18	64.3	152	4	US-08-672-850-12
5	18	64.3	200	4	US-09-149-476-495
6	18	64.3	237	2	US-08-933-750C-32
7	18	64.3	237	4	US-09-234-613-32
8	18	64.3	247	3	US-08-969-644-18
9	18	64.3	247	3	US-08-444-189-18
10	18	64.3	247	3	US-08-468-544-18
11	18	64.3	406	2	US-08-282-197C-52
12	18	64.3	420	2	US-08-282-197C-63
13	18	64.3	420	2	US-08-282-197C-66
14	18	64.3	451	2	US-08-468-812-8
15	18	64.3	451	4	US-08-590-563-8
16	18	64.3	705	2	US-08-663-566A-19
17	18	64.3	705	2	US-08-023-610-19
18	18	64.3	705	2	US-08-288-065A-19
19	18	64.3	705	2	US-08-362-240A-19
20	18	64.3	705	5	PCBUS-10245-19
21	18	64.3	705	5	PCBUS-10245-19
22	17	60.7	17	4	US-08-896-933-19
23	17	60.7	17	4	US-09-314-235-6
24	17	60.7	17	4	US-09-314-235-19
25	17	60.7	24	3	US-08-838-413A-20
26	17	60.7	24	3	US-08-838-413A-27
27	17	60.7	142	4	US-08-672-850-8

28	17	60.7	193	4	US-08-896-933-31	Sequence 31, Appl
29	17	60.7	193	4	US-09-314-235-31	Sequence 31, Appl
30	17	60.7	194	1	US-08-446-918A-6	Sequence 6, Appl
31	17	60.7	194	2	US-08-580-806-6	Sequence 6, Appl
32	17	60.7	234	1	US-08-152-456A-2	Sequence 2, Appl
33	17	60.7	234	1	US-08-440-221-2	Sequence 2, Appl
34	17	60.7	234	3	US-08-486-099-111	Sequence 111, App
35	17	60.7	234	3	US-08-360-107A-121	Sequence 121, App
36	17	60.7	234	3	US-08-484-223B-111	Sequence 111, App
37	17	60.7	234	3	US-08-919-597-111	Sequence 111, App
38	17	60.7	234	3	US-08-475-666A-111	Sequence 111, App
39	17	60.7	234	3	US-08-485-551A-111	Sequence 111, App
40	17	60.7	234	3	US-08-471-913A-111	Sequence 111, App
41	17	60.7	234	4	US-08-485-264A-111	Sequence 111, App
42	17	60.7	234	4	US-08-474-239A-111	Sequence 111, App
43	17	60.7	238	4	US-09-144-776B-12	Sequence 12, Appl
44	17	60.7	238	4	US-08-896-933-28	Sequence 28, Appl
45	17	60.7	238	4	US-09-314-235-28	Sequence 28, Appl

#### ALIGNMENTS

RESULT 1  
US-08-672-850-13  
; Sequence 13, Application US/08672850  
; Patent No. 6140117  
GENERAL INFORMATION:  
APPLICANT: Milbrandt, Jeffrey  
ATTORNEY/AGENT INFORMATION:  
NAME: Silva, Robin M.  
REGISTRATION NUMBER: 38,304  
REFERENCE/DOCKET NUMBER: A-63610  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 100 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-672-850-13

Query Match 64.3%; Score 18; DB 4; Length 100;  
Best Local Similarity 40.0%; Pred. No. 96;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 XXXXXXXXLD 12  
DB 44 KKSNAESMLD 53

RESULT 2  
US-08-672-850-4  
Sequence 4, Application US/08672850  
Patent No. 6140117  
GENERAL INFORMATION:  
APPLICANT: Milbrandt, Jeffrey  
APPLICANT: Araki, Toshiyuki  
TITLE OF INVENTION: NINTURIN  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
STREET: Four Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: United States  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/672,850  
FILING DATE: 24-JUL-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Silva, Robin M.  
REGISTRATION NUMBER: 38,304  
REFERENCE/DOCKET NUMBER: A-63610  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-672-850-4

Query Match 64.3%; Score 18; DB 4; Length 152;  
Best Local Similarity 40.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KKKXXXXXLD 12  
DB 44 KKSAAESMLD 53

RESULT 3  
US-08-672-850-6  
Sequence 6, Application US/08672850  
Patent No. 6140117  
GENERAL INFORMATION:  
APPLICANT: Milbrandt, Jeffrey  
APPLICANT: Araki, Toshiyuki  
TITLE OF INVENTION: NINTURIN  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
STREET: Four Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: United States  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/672,850

FILING DATE: 24-JUL-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Silva, Robin M.  
REGISTRATION NUMBER: 38,304  
REFERENCE/DOCKET NUMBER: A-63610  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-672-850-6

Query Match 64.3%; Score 18; DB 4; Length 152;  
Best Local Similarity 40.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KKKXXXXXLD 12  
DB 44 KKSAAESMLD 53

RESULT 4  
US-08-672-850-12  
Sequence 12, Application US/08672850  
Patent No. 6140117  
GENERAL INFORMATION:  
APPLICANT: Milbrandt, Jeffrey  
APPLICANT: Araki, Toshiyuki  
TITLE OF INVENTION: NINTURIN  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
STREET: Four Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: United States  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/672,850  
FILING DATE: 24-JUL-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Silva, Robin M.  
REGISTRATION NUMBER: 38,304  
REFERENCE/DOCKET NUMBER: A-63610  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-672-850-12

Query Match 64.3%; Score 18; DB 4; Length 152;  
Best Local Similarity 40.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KKKXXXXXLD 12

Db 44 KKSAAESMLD 53

RESULT 5  
US-09-149-476-495  
Sequence 495, Application US/09149476  
Patent No. 6420526  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P2002P1  
CURRENT APPLICATION NUMBER: US/09/149,476  
EARLIER FILING DATE: 1998-09-08  
EARLIER APPLICATION NUMBER: PCT/US98/04493  
EARLIER FILING DATE: 1998-03-06  
EARLIER APPLICATION NUMBER: 60/040,162  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,333  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/038,621  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,626  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,334  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,336  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,163  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/047,600  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,615  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,597  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,502  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,633  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,583  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,617  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,618  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,503  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,592  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,581  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,584  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,500  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,587  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,492  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,598  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,613  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,582  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,596  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,612  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,632  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,601

EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,580  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,568  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,314  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,569  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,311  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,671  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,313  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,672  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,315  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/056,886  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,877  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,889  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,893  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,630  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,878  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,662  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,872  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,882  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,637  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,903  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,888  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,879  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,880  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,894  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,911  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,636  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,874  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,910  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,864  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,631  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,845  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,892  
EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/057,761  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/047,595  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,599  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,588  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,585  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,586  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,590  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,594  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,589  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,593  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,614  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,578  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,576  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/047,501  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,670  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/056,632  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,664  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,876  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,881  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,909  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,875  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,862  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,887  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,908  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/057,650  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/056,884  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057,669  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/049,610  
EARLIER FILING DATE: 1997-06-13  
EARLIER APPLICATION NUMBER: 60/061,060  
EARLIER FILING DATE: 1997-10-02

Query Match 64.3%; Score 18; DB 4; Length 200;  
Best Local Similarity 40.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KXXXXXXD 12  
Db 151 KKAASHRAD 160

RESULT 6  
US-08-933-750C-32  
; Sequence 32, Application US/08933750C

Patent No. 5932442  
GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Bandman, Olga  
APPLICANT: Shah, Purvi  
APPLICANT: Au-Young, Janice  
APPLICANT: Yue, Henry  
APPLICANT: Guegler, Karl J.  
APPLICANT: Corley, Neil C.  
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES  
NUMBER OF SEQUENCES: 98  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: PASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/933,750C  
FILING DATE: September 23, 1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0356 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 237 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: BLADNOT03  
CLONE: 1602473  
US-08-933-750C-32

Query Match 64.3%; Score 18; DB 2; Length 237;  
Best Local Similarity 40.0%; Pred. No. 2e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KXXXXXXD 12  
Db 189 KKAASHRAD 198

RESULT 7  
US-09-234-613-32  
; Sequence 32, Application US/09234613  
; Patent No. 6132973  
; GENERAL INFORMATION:  
; APPLICANT: Lal, Preeti  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Bandman, Olga  
; APPLICANT: Shah, Purvi  
; APPLICANT: Au-Young, Janice  
; APPLICANT: Yue, Henry  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES

NUMBER OF SEQUENCES: 98  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/234,613  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/933,750  
FILING DATE: September 23, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0356 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 237 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: BLADNOT03  
CLONE: 1602473  
US-09-234-613-32

Query Match 64.38; Score 18; DB 4; Length 237;  
Best Local Similarity 40.08; Pred. No. 2e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KXXXXXXLD 12  
||  
||  
Db 189 KKAASHRALD 198

RESULT 8  
US-08-969-644-18  
Sequence 18, Application US/08969644  
Patent No. 6096519  
GENERAL INFORMATION:  
APPLICANT: Rattl, Giulio  
APPLICANT: Comanducci, Maurizio  
APPLICANT: Tecce, Mario F.  
APPLICANT: Giuliani, Marzia M.  
TITLE OF INVENTION: PCTD PLASMID ISOLATED FROM CHLAMYDIA  
TITLE OF INVENTION: TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY  
TITLE OF INVENTION: THEM, RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID  
TITLE OF INVENTION: GENES IN HETEROLOGOUS SYSTEMS, PREPARATION OF SAID  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH  
STREET: 301 N. Washington Street  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22046-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/969,644  
FILING DATE: 13-NOV-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/467,152  
FILING DATE:  
APPLICATION NUMBER: US/07/661,820  
FILING DATE:  
APPLICATION NUMBER: IT MI 91A000314  
FILING DATE: 07-FEB-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Svensson, Leonard R.  
REGISTRATION NUMBER: 30,330  
REFERENCE/DOCKET NUMBER: 1267-202P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-241-1300  
TELEFAX: 703-241-2848  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 247 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-969-644-18

Query Match 64.38; Score 18; DB 3; Length 247;  
Best Local Similarity 40.08; Pred. No. 2e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KXXXXXXLD 12  
||  
||  
Db 13 KKNQTASLD 22

RESULT 9  
US-08-444-189-18  
Sequence 18, Application US/08444189  
Patent No. 6110705  
GENERAL INFORMATION:  
APPLICANT: Rattl, Giulio  
APPLICANT: Comanducci, Maurizio  
APPLICANT: Tecce, Mario F.  
APPLICANT: Giuliani, Marzia M.  
TITLE OF INVENTION: PCTD PLASMID ISOLATED FROM CHLAMYDIA  
TITLE OF INVENTION: TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY  
TITLE OF INVENTION: THEM, RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID  
TITLE OF INVENTION: GENES IN HETEROLOGOUS SYSTEMS, PREPARATION OF SAID  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH  
STREET: 301 N. Washington Street  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22046-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/444,189  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/180,528  
FILING DATE:  
APPLICATION NUMBER: US/07/991,512  
FILING DATE:  
APPLICATION NUMBER: US/07/661,820

;; FILING DATE:;;  
;; APPLICATION NUMBER: IT MI 91A000314;;  
;; FILING DATE: 07-FEB-1991;;  
;; ATTORNEY/AGENT INFORMATION:;;  
;; NAME: Svensson, Leonard R.;;  
;; REGISTRATION NUMBER: 30,330;;  
;; TELECOMMUNICATION INFORMATION:;;  
;; TELEPHONE: 703-241-1300;;  
;; TELEFAX: 703-241-2848;;  
;; TELEX: 248345;;  
;; INFORMATION FOR SEQ ID NO: 18:;;  
;; SEQUENCE CHARACTERISTICS:;;  
;; LENGTH: 247 amino acids;;  
;; TYPE: amino acid;;  
;; TOPOLOGY: linear;;  
;; MOLECULE TYPE: protein;;  
US-08-444-189-18

Query Match 64.3%; Score 18; DB 3; Length 247;  
Best Local Similarity 40.0%; Pred. No. 2e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KKKXXXXXLD 12  
DB 13 KKNQTAASLD 22

RESULT 10  
US-08-468-544-18  
; Sequence 18, Application US/08468544  
; Patent No. 628563

;; GENERAL INFORMATION:;;  
;; APPLICANT: Ratti, Giulio;;  
;; APPLICANT: Comanducci, Maurizio;;  
;; APPLICANT: Tecce, Mario F.;;  
;; APPLICANT: Gullani, Marzia M.;;  
;; TITLE OF INVENTION: PCID PLASMID ISOLATED FROM CHLAMYDIA;;  
;; TITLE OF INVENTION: TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY;;  
;; TITLE OF INVENTION: THEM; RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID;;  
;; NUMBER OF SEQUENCES: 23;;  
;; CORRESPONDENCE ADDRESS:;;  
;; ADDRESSER: BIRCH, STEWART, KOLASCH & BIRCH;;  
;; STREET: 301 N. Washington Street;;  
;; CITY: Falls Church;;  
;; STATE: Virginia;;  
;; COUNTRY: USA;;  
;; ZIP: 22046-0747;;  
;; COMPUTER READABLE FORM:;;  
;; MEDIUM TYPE: Floppy disk;;  
;; COMPUTER: IBM PC compatible;;  
;; OPERATING SYSTEM: PC-DOS/MS-DOS;;  
;; SOFTWARE: Patentin Release #1.0, Version #1.25;;  
;; CURRENT APPLICATION DATA:;;  
;; APPLICATION NUMBER: US/08/468,544;;  
;; FILING DATE: 06-JUN-1995;;  
;; CLASSIFICATION: 435;;  
;; PRIOR APPLICATION DATA:;;  
;; APPLICATION NUMBER: US 07/661,820;;  
;; FILING DATE: 28-FEB-1991;;  
;; APPLICATION NUMBER: IT MI 91A000314;;  
;; ATTORNEY/AGENT INFORMATION:;;  
;; NAME: Svensson, Leonard R.;;  
;; REGISTRATION NUMBER: 30,330;;  
;; TELECOMMUNICATION INFORMATION:;;  
;; TELEPHONE: 703-241-1300;;  
;; TELEFAX: 703-241-2848;;  
;; TELEX: 248345;;  
;; INFORMATION FOR SEQ ID NO: 18:;;  
;; SEQUENCE CHARACTERISTICS:;;

;; LENGTH: 247 amino acids;;  
;; TYPE: amino acid;;  
;; TOPOLOGY: linear;;  
;; MOLECULE TYPE: protein;;  
US-08-468-544-18

Query Match 64.3%; Score 18; DB 4; Length 247;  
Best Local Similarity 40.0%; Pred. No. 2e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KKKXXXXXLD 12  
DB 13 KKNQTAASLD 22

RESULT 11  
US-08-282-197C-52  
; Sequence 52, Application US/08282197C  
; Patent No. 5871730

;; GENERAL INFORMATION:;;  
;; APPLICANT: Brzezinski, Ryszard;;  
;; APPLICANT: Dery, Claude V.;;  
;; APPLICANT: Beaulieu, Carole;;  
;; TITLE OF INVENTION: Thermostable Xylanase DNA, Protein and;;  
;; TITLE OF INVENTION: Methods of Use;;  
;; NUMBER OF SEQUENCES: 67;;  
;; CORRESPONDENCE ADDRESS:;;  
;; ADDRESSER: Sterne, Kessler, Goldstein & Fox P.L.L.C.;;  
;; STREET: 1100 New York Ave., NW;;  
;; CITY: Washington;;  
;; STATE: DC;;  
;; COUNTRY: USA;;  
;; ZIP: 20005;;  
;; COMPUTER READABLE FORM:;;  
;; MEDIUM TYPE: Floppy disk;;  
;; COMPUTER: IBM PC compatible;;  
;; OPERATING SYSTEM: PC-DOS/MS-DOS;;  
;; SOFTWARE: Patentin Release #1.0, Version #1.25;;  
;; CURRENT APPLICATION DATA:;;  
;; APPLICATION NUMBER: US/08/282,197C;;  
;; FILING DATE: 29-JUL-1994;;  
;; CLASSIFICATION: 435;;  
;; ATTORNEY/AGENT INFORMATION:;;  
;; NAME: Cimbalia, Michele A.;;  
;; REGISTRATION NUMBER: 33,851;;  
;; REFERENCE/DOCKET NUMBER: 1050.0410000;;  
;; TELECOMMUNICATION INFORMATION:;;  
;; TELEPHONE: 202-371-2600;;  
;; TELEFAX: 202-371-2540;;  
;; INFORMATION FOR SEQ ID NO: 52:;;  
;; SEQUENCE CHARACTERISTICS:;;  
;; LENGTH: 406 amino acids;;  
;; TYPE: amino acid;;  
;; TOPOLOGY: both;;  
US-08-282-197C-52

Query Match 64.3%; Score 18; DB 2; Length 406;  
Best Local Similarity 40.0%; Pred. No. 3.1e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KKKXXXXXLD 12  
DB 330 KKAAYTAVLD 339

RESULT 12  
US-08-282-197C-63  
; Sequence 63, Application US/08282197C  
; Patent No. 5871730  
;; GENERAL INFORMATION:;;  
;; APPLICANT: Brzezinski, Ryszard;;  
;; APPLICANT: Dery, Claude V.;;  
;; APPLICANT: Beaulieu, Carole;;



TITLE OF INVENTION: Thermostable Xylanase DNA, Protein and  
TITLE OF INVENTION: Methods of Use  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
STREET: 1100 New York Ave., NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/282,197C  
FILING DATE: 29-JUL-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Cimbalala, Michele A  
REGISTRATION NUMBER: 33,851  
REFERENCE/DOCKET NUMBER: 1050.0410000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 63:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 420 amino acids  
TYPE: amino acid  
TOPOLOGY: both  
US-08-282-197C-63

Query Match 64.3%; Score 18; DB 2; Length 420;  
Best Local Similarity 40.0%; Pred. No. 3.1e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KKKXXXXXD 12  
DB 330 KKAAYTAVLD 339

RESULT 13  
US-08-282-197C-66  
Sequence 66, Application US/08282197C  
Patent No. 5871730  
GENERAL INFORMATION:  
APPLICANT: Brzezinski, Ryszard  
APPLICANT: Dery, Claude V  
TITLE OF INVENTION: Thermostable Xylanase DNA, Protein and  
TITLE OF INVENTION: Methods of Use  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
STREET: 1100 New York Ave., NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/282,197C  
FILING DATE: 29-JUL-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Cimbalala, Michele A  
REGISTRATION NUMBER: 33,851  
REFERENCE/DOCKET NUMBER: 1050.0410000

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 66:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 420 amino acids  
TYPE: amino acid  
TOPOLOGY: both  
US-08-282-197C-66

Query Match 64.3%; Score 18; DB 2; Length 420;  
Best Local Similarity 40.0%; Pred. No. 3.1e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KKKXXXXXD 12  
DB 330 KKAAYTAVLD 339

RESULT 14  
US-08-468-812-8  
Sequence 8, Application US/08468812  
Patent No. 5935836  
GENERAL INFORMATION:  
APPLICANT: Vehmaenper, Jari  
APPLICANT: M ntyl, Arja  
APPLICANT: Fagerstr m, Richard  
APPLICANT: Lantto, Raija  
APPLICANT: Palohelmo, Marja  
APPLICANT: Suominen, Pirkko  
APPLICANT: Lahtinen, Taina  
APPLICANT: Kistio, Paula  
TITLE OF INVENTION: Actinomadura Xylanase Sequences and Methods  
TITLE OF INVENTION: of Use  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLSTEIN & FOX  
STREET: 1100 New York Ave., N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,812  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/332,412  
FILING DATE: 31-OCT-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/282,001  
FILING DATE: 29-JUL-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Bugalsky, Larry B.  
REGISTRATION NUMBER: 35,086  
REFERENCE/DOCKET NUMBER: 1050.0340002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 491 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

POSITION IN GENOME:  
CHROMOSOME/SEGMENT: M64551  
US-08-468-812-8

Query Match 64.3%; Score 18; DB 2; Length 491;  
Best Local Similarity 40.0%; Pred. No. 3.6e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KKKXXXXXID 12  
DB 333 KKAAYTAVLD 342

Query Match 64.3%; Score 18; DB 4; Length 491;  
Best Local Similarity 40.0%; Pred. No. 3.6e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
OY 3 KKKXXXXXID 12  
DB 333 KKAAYTAVLD 342  
Search completed: June 5, 2003, 15:32:44  
Job time : 10.5455 secs

## RESULT 15

US-08-590-563-8  
Sequence 8, Application US/08590563

Patent No. 6300114  
GENERAL INFORMATION:

APPLICANT: M uyl, Arja

APPLICANT: Vehmaaper, Jari

APPLICANT: Pagarstr m, Richard

APPLICANT: Lantto, Raija

APPLICANT: Palohelmo, Marja

APPLICANT: Suominen, Pirkko

APPLICANT: Lahtinen, Tarja

TITLE OF INVENTION: Production and Secretion of Proteins of

NUMBER OF SEQUENCES: 39

CORRESPONDENCE ADDRESS:

ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C.

STREET: 1100 New York Ave., N.W. Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/590,563

FILING DATE: 26-JAN-1996

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/332,412

FILING DATE: 31-OCT-1994

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/282,001

FILING DATE: 29-JUL-1994

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Bugalsky, Lawrence B.

REGISTRATION NUMBER: 35,086

REFERENCE/DOCKET NUMBER: 1050.0340003

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 491 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

POSITION IN GENOME:

CHROMOSOME/SEGMENT: M64551

US-08-590-563-8

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:31:01 ; Search time 14.5455 Seconds  
(without alignments)  
85.173 Million cell updates/sec

Title: US-09-150-947f-17  
Perfect score: 28  
Sequence: 1 XXXXXXXXXXXXLD 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues  
Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA.\*  
1: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubppa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubppa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubppa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubppa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubppa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubppa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubppa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubppa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubppa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	64.3	200	9	US-09-809-391-495 Sequence 495, App
2	18	64.3	237	10	US-09-840-787-32 Sequence 32, App1
3	18	64.3	365	9	US-10-125-692-16 Sequence 16, App1
4	18	64.3	491	10	US-09-770-621-8 Sequence 8, App1
5	17	60.7	12	9	US-10-113-809-3 Sequence 3, App1
6	17	60.7	58	12	US-10-001-879-145 Sequence 145, App
7	17	60.7	107	9	US-09-465-714-4 Sequence 4, App1
8	17	60.7	142	10	US-09-729-674-76 Sequence 76, App1
9	17	60.7	151	9	US-09-465-714-2 Sequence 2, App1
10	17	60.7	194	9	US-09-465-714-3 Sequence 3, App1
11	17	60.7	233	1	US-08-882-431-12 Sequence 12, App1
12	17	60.7	234	9	US-09-870-759-18 Sequence 18, App1
13	17	60.7	234	9	US-10-002-784-12 Sequence 12, App1
14	17	60.7	238	9	US-09-465-714-1 Sequence 1, App1
15	17	60.7	266	1	US-08-882-431-14 Sequence 14, App1
16	17	60.7	266	9	US-09-870-759-12 Sequence 12, App1
17	17	60.7	266	9	US-10-002-784A-14 Sequence 14, App1
18	17	60.7	383	9	US-09-738-626-6821 Sequence 6821, App
19	17	60.7	490	10	US-09-841-132-434 Sequence 434, App

20	17	60.7	582	9	US-10-046-938-29 Sequence 29, App1
21	17	60.7	623	9	US-09-738-626-5990 Sequence 5990, App
22	16	57.1	10	10	US-09-150-947B-3 Sequence 3, App1
23	16	57.1	10	10	US-09-150-947B-4 Sequence 4, App1
24	16	57.1	12	9	US-10-113-809-1 Sequence 1, App1
25	16	57.1	12	9	US-10-113-809-4 Sequence 4, App1
26	16	57.1	12	10	US-09-150-947B-2 Sequence 2, App1
27	16	57.1	13	10	US-09-150-947B-5 Sequence 5, App1
28	16	57.1	13	10	US-09-150-947B-6 Sequence 6, App1
29	16	57.1	13	10	US-09-150-947B-11 Sequence 11, App1
30	16	57.1	14	10	US-09-150-947B-9 Sequence 9, App1
31	16	57.1	14	10	US-09-150-947B-10 Sequence 10, App1
32	16	57.1	16	10	US-09-150-947B-7 Sequence 7, App1
33	16	57.1	25	10	US-09-803-165-29 Sequence 29, App1
34	16	57.1	26	10	US-09-150-947B-1 Sequence 1, App1
35	16	57.1	35	10	US-09-150-947B-8 Sequence 8, App1
36	16	57.1	36	9	US-09-764-891-4313 Sequence 4313, App
37	16	57.1	38	10	US-09-071-838-150 Sequence 150, App
38	16	57.1	39	10	US-09-843-845-12 Sequence 12, App1
39	16	57.1	63	9	US-10-044-692-65 Sequence 65, App1
40	16	57.1	63	9	US-10-044-539-65 Sequence 65, App1
41	16	57.1	69	9	US-09-843-676-16 Sequence 16, App1
42	16	57.1	69	9	US-09-766-253-16 Sequence 16, App1
43	16	57.1	69	9	US-09-438-486-16 Sequence 16, App1
44	16	57.1	69	9	US-10-053-758-16 Sequence 16, App1
45	16	57.1	69	9	US-10-054-295-16 Sequence 16, App1

## ALIGNMENTS

RESULT 1  
US-09-809-391-495  
Sequence 495, Application US/09809391  
Publication No. US20030049618A1  
GENERAL INFORMATION:  
APPLICANT: Ruben et al.  
TITLE OF INVENTION: 186 Human secreted proteins  
FILE REFERENCE: P2002P2  
CURRENT APPLICATION NUMBER: US/09/809,391  
CURRENT FILING DATE: 2001-03-16  
Prior application data removed - consult PAM or file wrapper  
NUMBER OF SEQ ID NOS: 761  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 495  
LENGTH: 200  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (200)  
OTHER INFORMATION: Xaa equals stop translation  
US-09-809-391-495

Query Match 64.3% Score 18; DB 9; Length 200;  
Best Local Similarity 40.0% Pred. No. 2e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KXXXXXXXXXLD 12  
DB 151 KKAASHRALD 160

RESULT 2  
US-09-840-787-32  
Sequence 32, Application US/09840787  
Patent No. US20020058264A1  
GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
Hillman, Jennifer L.  
Bandman, Olga  
Shah, Purvi  
Au-Young, Janice

Yue, Henry  
Guegler, Karl J.  
Corley, Neil C.  
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES  
NUMBER OF SEQUENCES: 98  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/840,787  
FILING DATE: 23-Apr-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/518,865  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0356 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 237 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: BLADNOT03  
CLONE: 1602473  
SEQUENCE DESCRIPTION: SEQ ID NO: 32 :  
US-09-840-787-32  
Query Match 64.3% Score 18; DB 10; Length 237;  
Best Local Similarity 40.0%; Pred. No. 2.3e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 3 KKKXXXXLTD 12  
DB 189 KKAASHRALD 198  
RESULT 3  
US-10-125-692-16  
Sequence 16, Application US/10125692  
Publication No. US20030044429a1  
GENERAL INFORMATION:  
APPLICANT: Aderem, Alan  
APPLICANT: Hayashi, Fumitaka  
APPLICANT: Smith, Kelly D.  
APPLICANT: Underhill, David M.  
TITLE OF INVENTION: Toll-Like Receptor 5 Ligands and Methods  
TITLE OF INVENTION: of use  
FILE REFERENCE: P-1S 5155  
CURRENT APPLICATION NUMBER: US/10/125,692  
CURRENT FILING DATE: 2002-04-17  
PRIOR APPLICATION NUMBER: US 60/285,477  
PRIOR FILING DATE: 2001-04-20  
NUMBER OF SEQ ID NOS: 43  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 16

LENGTH: 365  
TYPE: PRT  
ORGANISM: P. mirabilis1  
US-10-125-692-16  
Query Match 64.3% Score 18; DB 9; Length 365;  
Best Local Similarity 40.0%; Pred. No. 3.4e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 3 KKKXXXXLTD 12  
DB 197 KKNMAATLTD 206  
RESULT 4  
US-09-770-621-8  
Sequence 8, Application US/09770621  
Patent No. US20010024815A1  
GENERAL INFORMATION:  
APPLICANT: M ntyl, Arja  
APPLICANT: Vehmaaper, Jari  
APPLICANT: Fagerstr m, Richard  
APPLICANT: Lantto, Raija  
APPLICANT: Paloneimo, Marja  
APPLICANT: Suominen, Pirkko  
TITLE OF INVENTION: Production and Secretion of Proteins of  
NUMBER OF SEQUENCES: 39  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C.  
STREET: 1100 New York Ave., N.W. Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/770,621  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/590,563  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/332,412  
FILING DATE: 31-OCT-1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/282,001  
FILING DATE: 29-JUL-1994  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Bugalsky, Lawrence B.  
REGISTRATION NUMBER: 35,086  
REFERENCE/DOCKET NUMBER: 1050.0340003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2540  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 491 amino acids  
TYPE: amino acid  
STRANDEDNESS: No  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: M64551  
US-09-770-621-8

Query Match 64.3%; Score 18; DB 10; Length 491;  
Best Local Similarity 40.0%; Pred. No. 4.5e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KXXXXXXD 12  
DB 333 KKAAYTAVLD 342

RESULT 5  
US-10-113-809-3

Sequence 3, Application US/10113809  
Patent No. US2002017753A1  
GENERAL INFORMATION:  
APPLICANT: Maschke, Hans, E.  
TITLE OF INVENTION: EXOTOXIN-LIGAND  
FILE REFERENCE: MBP-011XX  
CURRENT APPLICATION NUMBER: US/10/113, 809  
CURRENT FILING DATE: 2002-03-29  
PRIOR APPLICATION NUMBER: 101 16 042.9-41  
PRIOR FILING DATE: 2001-03-30  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 12  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Bacterial Toxin Binding Ligand  
US-10-113-809-3

Query Match 60.7%; Score 17; DB 9; Length 12;  
Best Local Similarity 40.0%; Pred. No. 33;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KXXXXXXD 12  
DB 3 KQOLAISTLD 12

RESULT 6  
US-10-001-879-145  
Sequence 145, Application US/10001879  
Patent No. US20020127237A1  
GENERAL INFORMATION:  
APPLICANT: Salceda, Susana  
APPLICANT: Macina, Roberto  
APPLICANT: Reclon, Hervé  
APPLICANT: Caferkey, Robert  
APPLICANT: Ali, Shujath  
APPLICANT: Sun, Yongming  
APPLICANT: Liu, Chenghua  
TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and  
FILE REFERENCE: DEX-0281  
CURRENT APPLICATION NUMBER: US/10/001, 879  
CURRENT FILING DATE: 2001-11-20  
PRIOR APPLICATION NUMBER: 60/252,188  
PRIOR FILING DATE: 2000-11-21  
NUMBER OF SEQ ID NOS: 201  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 145  
LENGTH: 58  
TYPE: PRT  
ORGANISM: Homo sapien  
US-10-001-879-145

Query Match 60.7%; Score 17; DB 12; Length 58;  
Best Local Similarity 40.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KXXXXXXD 12  
DB 3 KXXXXXXD 12

DB 42 KKRRTQATLD 51

RESULT 7  
US-09-465-714-4

Sequence 76, Application US/09465714  
Publication No. US20030032582A1  
GENERAL INFORMATION:  
APPLICANT: Wahlsten, Jennifer L.  
APPLICANT: Ramakrishnan, S.  
APPLICANT: Schlievert, Patrick M.  
TITLE OF INVENTION: SPONTANEOUS MEMBRANE INSERTION  
FILE REFERENCE: 09531/003001  
CURRENT APPLICATION NUMBER: US/09/465, 714  
CURRENT FILING DATE: 1999-12-17  
PRIOR APPLICATION NUMBER: 09/001,593  
PRIOR FILING DATE: 1998-12-31  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 4  
LENGTH: 107  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Truncated bacterial sequence  
US-09-465-714-4

Query Match 60.7%; Score 17; DB 9; Length 107;  
Best Local Similarity 40.0%; Pred. No. 2.4e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KXXXXXXD 12  
DB 34 KQOLAISTLD 43

RESULT 8  
US-09-729-674-76  
Sequence 76, Application US/09729674  
Patent No. US20010039335A1  
GENERAL INFORMATION:  
APPLICANT: Jacobs, Kenneth  
APPLICANT: McCoy, John M.  
APPLICANT: Lavaille, Edward R.  
APPLICANT: Collins-Racie, Lisa A.  
APPLICANT: Evans, Cheryl  
APPLICANT: Werberg, David  
APPLICANT: Treacy, Maurice  
APPLICANT: Agostino, Michael J.  
APPLICANT: Steindinger II, Robert J.  
APPLICANT: Spaulding, Vikki  
APPLICANT: Wong, Gordon G.  
APPLICANT: Clark, Hilary  
APPLICANT: Rechelet, Kim  
APPLICANT: Genetics Institute, Inc.  
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM  
FILE REFERENCE: 6055-64X  
CURRENT APPLICATION NUMBER: US/09/729, 674  
CURRENT FILING DATE: 2000-12-04  
PRIOR APPLICATION NUMBER: 09/539, 330  
PRIOR FILING DATE: 2000-03-30  
NUMBER OF SEQ ID NOS: 283  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 76  
LENGTH: 142  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-729-674-76

Query Match 60.7%; Score 17; DB 10; Length 142;  
Best Local Similarity 40.0%; Pred. No. 3.1e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KXXXXXXLD 12  
DB 30 KKSVAESMLD 39

## RESULT 9

US-09-465-714-2  
; Sequence 2, Application US/09465714  
; Publication No. US20030032582A1  
; GENERAL INFORMATION:  
; APPLICANT: Wahlsten, Jennifer L.  
; APPLICANT: Ramakrishnan, S.  
; APPLICANT: Schlievert, Patrick M.  
; TITLE OF INVENTION: SPONTANEOUS MEMBRANE INSERTION  
; FILE REFERENCE: 09531/003001  
; CURRENT APPLICATION NUMBER: US/09/465,714  
; CURRENT FILING DATE: 1999-12-17  
; PRIOR APPLICATION NUMBER: 09/001,593  
; PRIOR FILING DATE: 1998-12-31  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 151  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Hybrid containing bacterial and mammalian sequence  
US-09-465-714-2

Query Match 60.7%; Score 17; DB 9; Length 151;  
Best Local Similarity 40.0%; Pred. No. 3.3e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KXXXXXXLD 12  
DB 34 KKOALAISTLD 43

## RESULT 10

US-09-465-714-3  
; Sequence 3, Application US/09465714  
; Publication No. US20030032582A1  
; GENERAL INFORMATION:  
; APPLICANT: Wahlsten, Jennifer L.  
; APPLICANT: Ramakrishnan, S.  
; APPLICANT: Schlievert, Patrick M.  
; TITLE OF INVENTION: SPONTANEOUS MEMBRANE INSERTION  
; FILE REFERENCE: 09531/003001  
; CURRENT APPLICATION NUMBER: US/09/465,714  
; CURRENT FILING DATE: 1999-12-17  
; PRIOR APPLICATION NUMBER: 09/001,593  
; PRIOR FILING DATE: 1998-12-31  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 194  
; TYPE: PRT  
; ORGANISM: Bacterial Sequence  
US-09-465-714-3

Query Match 60.7%; Score 17; DB 9; Length 194;  
Best Local Similarity 40.0%; Pred. No. 4.1e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KXXXXXXLD 12  
DB 121 KKOALAISTLD 130

RESULT 11  
US-08-882-431-12  
; Sequence 12, Application US/08882431  
; Publication No. US20030009015A1

GENERAL INFORMATION:  
APPLICANT: Robert G. Ulrich,  
APPLICANT: Mark A. Olson  
APPLICANT: Sina Bavari  
TITLE OF INVENTION: Bacterial Superantigen  
TITLE OF INVENTION: Vaccines  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John Moran  
STREET: US Army MCMC -504 Scott Street MCMR-JA (John Moran-Patent Atty)  
CITY: FORT DERRICK  
STATE: MARYLAND  
COUNTRY: USA  
ZIP: 21702-5012

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh 7.5  
SOFTWARE: Microsoft Word 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/882,431  
FILING DATE: June 25, 1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Moran, John  
REGISTRATION NUMBER: 26,313  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 619-2065  
TELEFAX: (301) 619-7714  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 233  
TYPE: Amino Acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
US-08-882-431-12

Query Match 60.7%; Score 17; DB 1; Length 233;  
Best Local Similarity 40.0%; Pred. No. 4.9e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KXXXXXXLD 12  
DB 160 KKOALAISTLD 169

## RESULT 12

US-09-870-759-18  
; Sequence 18, Application US/09870759  
; Patent No. US20020177551A1  
; GENERAL INFORMATION:  
; APPLICANT: TERMAN, David S  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE  
; FILE REFERENCE: 870759  
; CURRENT APPLICATION NUMBER: US/09/870,759  
; CURRENT FILING DATE: 2002-01-14  
; PRIOR APPLICATION NUMBER: US 60/208,128  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 166  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 18  
; LENGTH: 234  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-09-870-759-18

Query Match 60.7%; Score 17; DB 9; Length 234;  
Best Local Similarity 40.0%; Pred. No. 4.9e+02;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KXXXXXXLD 12  
11  
11

Db 161 KKOLAISTLD 170

RESULT 13  
US-10-002-784A-12  
; Sequence 12, Application US/10002784A  
; Publication No. US20030036644A1  
; GENERAL INFORMATION:  
; /33  
; APPLICANT: Ulrich, Robert G.  
; TITLE OF INVENTION: Bacterial Superantigen Vaccines  
; FILE REFERENCE: 003/233/SAP  
; CURRENT APPLICATION NUMBER: US/10/002,784A  
; CURRENT FILING DATE: 2001-11-26  
; PRIOR APPLICATION NUMBER: 08/882,431; 09/144,776  
; PRIOR FILING DATE: 97-06-25; 98-09-01  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: Apple Macintosh Microsoft Word 6.0  
; SEQ ID NO 12  
; LENGTH: 234  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: toxin shock syndrom toxin-1 mutant  
US-10-002-784A-12

Query Match 60.7%; Score 17; DB 9; Length 234;  
Best Local Similarity 40.0%; Pred. No. 4.9e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KXXXXXXLD 12  
11  
11

Db 161 KKOLAISTLD 170

RESULT 14  
US-09-465-714-1  
; Sequence 1, Application US/09465714  
; Publication No. US20030032582A1  
; GENERAL INFORMATION:  
; APPLICANT: Mahliken, Jennifer L.  
; APPLICANT: Ramakrishnan, S.  
; APPLICANT: Schlievert, Patrick M.  
; TITLE OF INVENTION: SPONTANEOUS MEMBRANE INSERTION  
; FILE REFERENCE: 09531/003001  
; CURRENT APPLICATION NUMBER: US/09/465,714  
; CURRENT FILING DATE: 1999-12-17  
; PRIOR APPLICATION NUMBER: 09/001,593  
; PRIOR FILING DATE: 1998-12-31  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 238  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Hybrid containing bacterial and mammalian sequence  
US-09-465-714-1

Query Match 60.7%; Score 17; DB 9; Length 238;  
Best Local Similarity 40.0%; Pred. No. 5e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KXXXXXXLD 12  
11  
11

Db 121 KKOLAISTLD 130

RESULT 15

US-08-882-431-14  
; Sequence 14, Application US/0882431  
; Publication No. US20030009015A1  
; GENERAL INFORMATION:  
; APPLICANT: Robert G. Ulrich,  
; APPLICANT: Mark A. Olson  
; APPLICANT: Sina Bavari  
; TITLE OF INVENTION: Bacterial Superantigen  
; TITLE OF INVENTION: Vaccines  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: John Moran  
; STREET: US Army MRC -504 Scott Street MCMR-JA (John Moran-Patent Atty)  
; CITY: FORT DETRICK  
; STATE: MARYLAND  
; COUNTRY: USA  
; ZIP: 21702-5012  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Macintosh 7.5  
; SOFTWARE: Microsoft Word 6.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/882,431  
; FILING DATE: June 25, 1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Moran, John  
; REGISTRATION NUMBER: 26,313  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 619-7714  
; TELEFAX: (301) 619-2065  
; INFORMATION FOR SEQ ID NO: 14:  
; . SEQUENCE CHARACTERISTICS:  
; LENGTH: 266  
; TYPE: Amino Acid  
; STRANDEDNESS: Unknown  
; TOPOLOGY: Unknown  
; MOLECULE TYPE: Peptide  
US-08-882-431-14

Query Match 60.7%; Score 17; DB 1; Length 266;  
Best Local Similarity 40.0%; Pred. No. 5.5e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KXXXXXXLD 12  
11  
11

Db 178 KSVTAQELD 187

Search completed: June 5, 2003, 15:49:05  
Job time: 14.5455 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: June 5, 2003, 15:24:01 ; Search time 11.8162 Seconds  
(without alignments)  
97.613 Million cell updates/sec

Title: US-09-150-947f-17  
Perfect score: 28  
Sequence: 1 XXXXXXXXXXXX 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	67.9	284	2	conserved hypothet
2	19	67.9	378	2	probable elongatio
3	18	64.3	100	1	ribosomal protein
4	18	64.3	159	2	hypothetical prote
5	18	64.3	161	2	transcription regu
6	18	64.3	169	2	hypothetical prote
7	18	64.3	171	2	hypothetical prote
8	18	64.3	212	2	probable tet opero
9	18	64.3	212	2	probable tet opero
10	18	64.3	212	2	probable transcrip
11	18	64.3	217	2	hypothetical prote
12	18	64.3	237	2	hypothetical prote
13	18	64.3	247	2	hypothetical prote
14	18	64.3	247	2	hypothetical prote
15	18	64.3	278	2	hypothetical prote
16	18	64.3	278	2	hypothetical prote
17	18	64.3	344	2	probable portal pr
18	18	64.3	355	2	flagellin flhC-1-
19	18	64.3	383	2	MG349 homolog G12
20	18	64.3	413	2	actin like protein
21	18	64.3	427	2	endo-1,4-beta-xyla
22	18	64.3	477	1	hypothetical prote
23	18	64.3	533	2	hypothetical prote
24	18	64.3	837	2	hypothetical prote
25	18	64.3	858	2	probable permease,
26	18	64.3	863	2	hypothetical prote
27	18	64.3	1035	2	hypothetical prote
28	18	64.3	1042	2	hypothetical prote
29	18	64.3	1279	2	protoporphylin IX

30	17	60.7	109	2	A69448	hypothetical prote
31	17	60.7	118	2	T15655	hypothetical prote
32	17	60.7	137	2	G71541	probable camp-depe
33	17	60.7	137	2	C81695	cyclic nucleotide-
34	17	60.7	156	2	S23357	H+-transporting tw
35	17	60.7	185	2	G64084	3-methyladenine DN
36	17	60.7	186	2	S66070	conserved hypothet
37	17	60.7	207	2	AC2950	transcription regu
38	17	60.7	213	2	H98332	probable transcrip
39	17	60.7	217	1	H84322	heat shock protein
40	17	60.7	223	2	H84977	ribose-5-phosphate
41	17	60.7	226	2	A86268	hypothetical prote
42	17	60.7	234	1	XCSAS1	toxic shock syndro
43	17	60.7	234	2	B89992	toxic shock syndro
44	17	60.7	234	2	C69519	conserved hypothet
45	17	60.7	239	2	A83953	uridylylate kinase s

## ALIGNMENTS

RESULT 1  
T50261  
conserved hypothetical protein SPAC890.05 with G-patch domain [imported] - fission ye  
C/Species: Schizosaccharomyces pombe  
C/Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 09-Jun-2000  
C/Accession: T50261  
R/Saunders, D.; Harris, D.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, December 1999  
A/Reference number: Z25052  
A/Accession: T50261  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-284 <SNU>  
A/Cross-references: EMBL:AL133498; PIDN:CA863496.1; GSPDB:GN00066; SPDB:SPAC890.05  
A/Experimental source: strain 972h(-); cosmid c890  
C/Genetics:  
A/Gene: SPDB:SPAC890.05  
A/Map position: 1

Query Match 67.9% Score 19; DB 2; Length 284;  
Best Local Similarity 40.0%; Pred. No. 6.8;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KXXXXXXLD 12  
DB 179 KKTGTGSLD 188

RESULT 2  
E84806  
probable elongation factor [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C/Accession: E84806  
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.  
M.; Koo, H.; Moffat, K.S.; Croft, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,  
eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter  
Nature 402:761-768, 1999  
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A/Reference number: A84420; MUID:20083487; PMID:10617197  
A/Accession: E84806  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-378 <STO>  
A/Cross-references: GB:AE002093; MID:93786016; PIDN:AAC67362.1; GSPDB:GN00139  
C/Genetics:  
A/Gene: AT2G38560  
A/Map position: 2

Query Match 67.9% Score 19; DB 2; Length 378;  
Best Local Similarity 40.0%; Pred. No. 8.8;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KXXXXXXXXL 12  
||  
DB 13 KKAADAAALD 22

## RESULT 3

R3LV14

ribosomal protein S14, chloroplast - liverwort (Marchantia polymorpha) chloroplast  
C:Species: chloroplast Marchantia polymorpha  
C:Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 22-Jun-1999  
A:Accession: A02733; S01876

R:Ohya, K.  
submitted to the EMBL Data Library, October 1986  
A:Reference number: A00150

A:Accession: A02733  
A:Molecule type: DNA

A:Residues: 1-100 <OHY>

R:Ohya, K.; Fukuzawa, H.; Kohchi, T.; Shirai, H.; Sano, T.; Sano, S.; Umesono, K.; Shi  
Nature 322, 572-574, 1986

A:Title: Chloroplast gene organization deduced from complete sequence of liverwort March  
A:Reference number: A38014

A:Contents: annotation; gene organization, sites, features  
R:Umesono, K.; Inokuchi, H.; Shiki, Y.; Takeuchi, M.; Chang, Z.; Fukuzawa, H.; Kohchi, T

J. Mol. Biol. 203, 299-331, 1988  
A:Title: Structure and organization of Marchantia polymorpha chloroplast genome. II. Gen  
A:Reference number: S01567; MUID:89068686; PMID:2974085

A:Accession: S01876  
A:Molecule type: DNA

A:Residues: 1-100 <UME>

A:Cross-references: GB:X04465; GB:Y00686; NID:g11640; PIDN:CAA28083.1; PID:g11670  
C:Genetics:

A:Gene: rps14  
A:Genome: chloroplast

C:Superfamily: Escherichia coli ribosomal protein S14  
C:Keywords: chloroplast; protein biosynthesis; ribosome

Query Match 64.3%; Score 18; DB 1; Length 100;  
Best Local Similarity 40.0%; Pred. No. 7.8;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KXXXXXXXXL 12  
||  
DB 29 KKTETSSLD 38

## RESULT 4

T48837

hypothetical protein 68B2.210 [imported] - Neurospora crassa  
C:Species: Neurospora crassa  
C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 16-Feb-2001

C:Accession: T48837  
R:Schulte, U.; Altm, V.; Hohnsels, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,  
submitted to the Protein Sequence Database, April 2000  
A:Reference number: Z24541

A:Accession: T48837  
A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-159 <SCH>

A:Cross-references: EMBL:AL353821; GSPDB:GN00112; NCSP:68B2.210  
A:Experimental source: cosmid contig 68B2, strain 74

C:Genetics:

A:Gene: NCSP:68B2.210  
A:Map position: 2

C:Superfamily: Neurospora crassa hypothetical protein 68B2.210

Query Match 64.3%; Score 18; DB 2; Length 159;  
Best Local Similarity 40.0%; Pred. No. 12;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KXXXXXXXXL 12  
||  
DB 71 KKTALASLD 80

RESULT 5  
AB2401  
transcription regulator Tetr family VCA0917 [imported] - Vibrio cholerae (strain N169

C:Species: Vibrio cholerae  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001

C:Accession: AB2401  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.

Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers  
L. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: AB2035; MUID:20406833; PMID:10952301

A:Accession: AB2401  
A:Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-161 <HEI>

A:Cross-references: GB:AE004420; GB:AE003853; NID:g9658361; PIDN:AAF96814.1; GSPDB:GN  
A:Experimental source: serogroup O1, strain N16961; biotype El Tor

C:Genetics:

A:Gene: VCA0917  
A:Map position: 2

Query Match 64.3%; Score 18; DB 2; Length 161;  
Best Local Similarity 40.0%; Pred. No. 12;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KXXXXXXXXL 12  
||  
DB 62 KKTDFASALD 71

## RESULT 6

S07748

hypothetical protein 13 - Paramecium tetraurelia mitochondrion  
C:Species: mitochondrion Paramecium tetraurelia

C:Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 07-Dec-1999  
C:Accession: S07748

R:Pritchard, A.E.; Seilhammer, J.J.; Mahalingam, R.; Sable, C.L.; Venutti, S.E.; Cummin  
Nucleic Acids Res. 18, 173-180, 1990

A:Title: Nucleotide sequence of the mitochondrial genome of Paramecium.  
A:Reference number: S07725; MUID:90174913; PMID:2308823

A:Accession: S07748  
A:Status: translation not shown

A:Molecule type: DNA  
A:Residues: 1-169 <PRI>

A:Cross-references: EMBL:X15917; NID:g13256; PID:g578763  
C:Genetics:

A:Genome: mitochondrion  
A:Genetic code: SGC6

A:Start codon: ATA  
C:Superfamily: Paramecium tetraurelia mitochondrion hypothetical protein 13  
C:Keywords: mitochondrion

Query Match 64.3%; Score 18; DB 2; Length 169;  
Best Local Similarity 40.0%; Pred. No. 13;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KXXXXXXXXL 12  
||  
DB 36 KKTSTYSKLD 45

## RESULT 7

A96723

hypothetical protein F20P5.23 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: A96723

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federle, N.A.; Kaul, S.; White, O.; Alon  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,  
ansen, N.F.; Hughes, B.; Huizart, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marziani, R.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, K.M.; Yu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: A96723

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-171 <STO>

A:Cross-references: GB:AE005173; NID:92194134; PIDN:AAB61109.1; GSPDB:GN00141

C:Genetics:

A:Gene: F20P5.23

A:Map position: 1

Query Match

Best Local Similarity 64.3%; Score 18; DB 2; Length 171;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KKKXXXXXXLD 12

Db 16 KKLTITDAMD 25

RESULT 8

C90786

probable tet operon regulator [imported] - *Escherichia coli* (strain O157:H7, substrain H

C:Species: *Escherichia coli*

C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001

C:Accession: C90786

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; gasawa, N.; Yasunaga, T.; Kuahara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gene

A:Reference number: A9629; MUID:21136231; PMID:11258796

A:Accession: C90786

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-212 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA834682.1; PID:q13360719; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RMD 0509952

C:Genetics:

A:Gene: ECS1259

Query Match

Best Local Similarity 64.3%; Score 18; DB 2; Length 212;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KKKXXXXXXLD 12

Db 20 KKAIIISALD 29

RESULT 9

A85646

probable tet operon regulator ydcC [imported] - *Escherichia coli* (strain O157:H7, substr

C:Species: *Escherichia coli*

C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001

C:Accession: A85646

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: A85646

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-212 <STO>

A:Cross-references: GB:AE005174; NID:912514373; PIDN:AAG55629.1; GSPDB:GN00145; OMGP:Z15

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: ydcC

Query Match

Best Local Similarity 64.3%; Score 18; DB 2; Length 212;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KKKXXXXXXLD 12

Db 20 KKAIIISALD 29

RESULT 10

C64843

probable transcription regulator ydcC - *Escherichia coli* (strain K-12)

C:Species: *Escherichia coli*

C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002

C:Accession: C64843

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of *Escherichia coli* K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: C64843

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-212 <BLAT>

A:Cross-references: GB:AE000203; GB:U00096; NID:q1787248; PIDN:AAC74098.1; PID:q17872

A:Experimental source: strain K-12, substrain MO1655

C:Genetics:

A:Gene: ydcC

A:Keywords: DNA binding; transcription regulation

F:39-58/Region: helix-turn-helix motif

Query Match

Best Local Similarity 64.3%; Score 18; DB 2; Length 212;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KKKXXXXXXLD 12

Db 20 KKAIIISALD 29

RESULT 11

B64080

hypothetical protein HI0601 - *Haemophilus influenzae* (strain Rd KW20)

C:Species: *Haemophilus influenzae*

C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 08-Oct-1999

C:Accession: B64080

R:Feilschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Feldman, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995

A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente A:Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.

A:Reference number: A64000; MUID:95350630; PMID:7542800

A:Accession: B64080

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-217 <TIGR>

A:Cross-references: GB:U32741; GB:I42023; NID:q1573582; PIDN:AAC22258.1; PID:q1573592

Query Match

Best Local Similarity 64.3%; Score 18; DB 2; Length 217;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KKKXXXXXXLD 12

Db 161 KKSAGATLD 170

RESULT 12

T14770

hypothetical protein DKFZP566E144.1 - human

C:Species: Homo sapiens (man)  
 C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 11-Jan-2000  
 C:Accession: T14770  
 R:Boecker, H.; Boecker, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
 submitted to the Protein Sequence Database, August 1999  
 A:Reference number: 218182  
 A:Accession: T14770  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-237 <BL0>  
 A:Cross-References: EMBL:AL110239  
 A:Experimental source: fetal kidney; clone DKFZp566E144  
 C:Genetics:  
 A>Note: DKFZp566E144.1  
 C:Superfamily: human conserved hypothetical protein DKFZp566E144.1

Query Match 64.3% Score 18; DB 2; Length 237;  
 Best Local Similarity 40.0%; Pred. No. 17;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KKKXXXXXXLD 12  
 || ||  
 DB 189 KKAASHRALD 198

## RESULT 13

S01924  
 hypothetical protein 6 - Chlamydia trachomatis plasmids  
 N:Alternate names: hypothetical protein P-10  
 C:Species: Chlamydia trachomatis  
 C>Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 08-Oct-1999  
 C:Accession: S01924; S00794  
 R:Comanducci, M.; Ricci, S.; Ratti, G.  
 M.O. Microbiol. 2, 531-538, 1988

A:Title: The structure of a plasmid of Chlamydia trachomatis believed to be required for  
 A:Reference number: S01920; MUID:89013895; PMID:2845228  
 A:Accession: S01924  
 A:Molecule type: DNA  
 A:Residues: 1-247 <CON>  
 A:Cross-References: EMBL:X07547; NID:940730; PIDN:CAA30426.1; PID:9581014  
 A:Experimental source: plasmid pCHL1  
 R:Hate, C.; Ward, M.E.; Clarke, I.N.  
 A>Note: the authors translated the initiation codon GTG for residue 1 as Val  
 Nucleic Acids Res. 16, 4053-4067, 1988  
 A:Title: Analysis of the entire nucleotide sequence of the cryptic plasmid of Chlamydia  
 A:Reference number: S00788; MUID:88233398; PMID:2836808  
 A:Accession: S00794  
 A:Molecule type: DNA  
 A:Residues: 175-247 <HAT>  
 A:Cross-References: EMBL:X06707  
 A:Experimental source: plasmid pLGV440  
 C:Genetics:  
 A:Genome: plasmid  
 A:Start codon: GTG

Query Match 64.3% Score 18; DB 2; Length 247;  
 Best Local Similarity 40.0%; Pred. No. 18;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KKKXXXXXXLD 12  
 || ||  
 DB 13 KKNOTASLD 22

## RESULT 14

F37386  
 hypothetical protein 6 - Chlamydia trachomatis (serotype D) plasmid pCHL1  
 C:Species: Chlamydia trachomatis  
 C>Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 17-Feb-1994  
 C:Accession: F37386  
 R:Comanducci, M.; Ricci, S.; Cevenini, R.; Ratti, G.  
 plasmid 23, 149-154, 1990  
 A:Title: Diversity of the Chlamydia trachomatis common plasmid in biovars with different

A:Reference number: A37386; MUID:90301796; PMID:2194229  
 A:Accession: F37386  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-247 <CON>  
 A:Cross-References: GB:J03321  
 C:Genetics:  
 A:Genome: plasmid

Query Match 64.3% Score 18; DB 2; Length 247;  
 Best Local Similarity 40.0%; Pred. No. 18;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KKKXXXXXXLD 12  
 || ||  
 DB 13 KKNOTASLD 22

## RESULT 15

AB3091  
 hypothetical protein Atu4350 [imported] - Agrobacterium tumefaciens (strain C58, Dupo  
 C:Species: Agrobacterium tumefaciens  
 C>Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
 C:Accession: AB3091  
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo  
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuyavain, T.; Levy, R.; Li, M.; McCl  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam  
 ster, E.W.  
 A:Title: The genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A:Reference number: AB2577; PMID:11743193  
 A:Accession: AB3091  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-278 <KUP>  
 A:Cross-References: GB:AE008689; PIDN:AAL45144.1; PID:917742817; GSPDB:GN00187  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: Atu4350  
 A:Map position: linear chromosome

Query Match 64.3% Score 18; DB 2; Length 278;  
 Best Local Similarity 40.0%; Pred. No. 20;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KKKXXXXXXLD 12  
 || ||  
 DB 184 KKLATHTLD 193

Search completed: June 5, 2003, 15:34:01  
 Job time : 12.8182 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:34:07 ; Search time 7.81818 Seconds  
(without alignments)  
63.661 Million cell updates/sec

Title: US-09-150-947f-17  
Perfect score: 28  
Sequence: 1 XXXXXXXXXXXXLD 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues  
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	18	64.3	100 1	RRL4_MARPO
2	18	64.3	152 1	NIJ1_HUMAN
3	18	64.3	152 1	NIJ1_MOUSE
4	18	64.3	152 1	NIJ1_RAT
5	18	64.3	169 1	YML3_PASTE
6	18	64.3	212 1	YCDC_ECOLI
7	18	64.3	217 1	TPOX_HAEIN
8	18	64.3	237 1	ORN_HUMAN
9	18	64.3	247 1	GP6D_CHLTR
10	18	64.3	344 1	FLC1_PROMI
11	18	64.3	365 1	Y349_MYCPN
12	18	64.3	413 1	ARP3_SCHPO
13	18	64.3	427 1	ARP3_SCHPO
14	18	64.3	477 1	XRNA_STRLI
15	18	64.3	837 1	YL34_YEAST
16	18	64.3	1279 1	BCNR_CHLVI
17	17	60.7	81 1	EX7S_PASMU
18	17	60.7	118 1	YQ93_CAEEL
19	17	60.7	142 1	NIJ2_HUMAN
20	17	60.7	143 1	NIJ2_MOUSE
21	17	60.7	144 1	NIJ2_RAT
22	17	60.7	147 1	RLIS_MYCGA
23	17	60.7	156 1	ATPX_OPOSI
24	17	60.7	185 1	3MGA_HAEIN
25	17	60.7	186 1	YABF_BACSU
26	17	60.7	211 1	ORN_DROME
27	17	60.7	217 1	GRPE_MYCGE
28	17	60.7	223 1	GRPE_MYCGE
29	17	60.7	234 1	TSST_STPAU
30	17	60.7	256 1	YB32_MYCPN
31	17	60.7	266 1	ETC1_STPAU
32	17	60.7	266 1	ETC2_STPAU
33	17	60.7	266 1	ETC3_STPAU

34	17	60.7	272 1	TR05_CAMJE	Q9P12 campylobact
35	17	60.7	276 1	MYO1_ONCMY	Q91205 oncorhynch
36	17	60.7	307 1	ERA_NEIMA	Q91422 neisaria m
37	17	60.7	307 1	ERA_NEIMA	Q91422 neisaria m
38	17	60.7	347 1	YAB9_MYCPN	P75593 mycoplasma
39	17	60.7	363 1	LE3A_ASFNG	P87236 aspergillus
40	17	60.7	368 1	DXR_HELPJ	Q9ZML6 helicobacte
41	17	60.7	390 1	CNX2_ARATH	Q39055 arabidopsis
42	17	60.7	399 1	Y199_ARCFU	Q28380 archaeglob
43	17	60.7	413 1	YB44_MYCPN	P75142 mycoplasma
44	17	60.7	457 1	PRC1_STRGR	P52320 streptomyce
45	17	60.7	467 1	LE02_ANASP	Q8YX02 anabaena sp

## ALIGNMENTS

RESULT 1  
RRL4\_MARPO STANDARD; PRT; 100 AA.  
ID AC P06370;  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Chloroplast 30S ribosomal protein S14.  
GN RPS14.  
OS Marchantia polymorpha (Liverwort).  
OC Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Marchantiophyta;  
OC Marchantiopsida; Marchantiales; Marchantiaaceae; Marchantia.  
OC Marchantiaceae; Marchantia.  
OX NCBI\_TaxID=3197;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85087956; PubMed=6393057;  
RA Umesono K., Inokuchi H., Ohyama K., Ozeki H.;  
RT "Nucleotide sequence of Marchantia polymorpha chloroplast DNA: a  
RT region possibly encoding three tRNAs and three proteins including a  
RT homologue of E. coli ribosomal protein S14.";  
RL Nucleic Acids Res. 12:9551-9565(1984).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Ohyama K., Fukuzawa H., Kohchi T., Shirai H., Sano T., Sano S.,  
RA Umesono K., Shiki Y., Takeuchi M., Chang Z., Aota S., Inokuchi H.,  
RA Ozeki H.;  
RT "Chloroplast gene organization deduced from complete sequence of  
RT liverwort Marchantia polymorpha chloroplast DNA.";  
RL Nucleic Acids Res. 12:9551-9565(1984).  
CC -1- SIMILARITY: BELONGS TO THE S14P FAMILY OF RIBOSOMAL PROTEINS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
DR EMBL: X01647; CA25805.1; -;  
DR EMBL: X04465; CA28083.1; -;  
DR PIR: A02733; R3LV14.  
DR InterPro: IPR001209; Ribosomal\_S14.  
DR Pfam: PF00253; Ribosomal\_S14; 1.  
DR PROSITE: PS00527; RIBOSOMAL\_S14; 1.  
KW Ribosomal protein; Chloroplast.  
SQ SEQUENCE 100 AA; 11880 MW; 5AB14F480963EEB0 CRC64;  
Query Match 64.3%; Score 18; DB 1; Length 100;  
Best Local Similarity 40.0%; Pred. No. 4.4;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
3 XXXXXXXXXXXXLD 12  
||  
||

DB 29 KRITESSID 38

RESULT 2

NIJL\_HUMAN STANDARD: PRT: 152 AA.

AC 09282: 09282: 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DE 16-OCT-2001 (Rel. 40, Last annotation update)

DE NINJ1 (Nerve injury-induced protein 1).

GN NINJ1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_TaxID=9606;

ON [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-96374367: PubMed-8780658;

RA Arai T., Milderand J.;

RT "Ninjurin, a novel adhesion molecule, is induced by nerve injury and promotes axonal growth."

RT Neuron 17:353-361(1996).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE-98126436: PubMed-9465296;

RA Chadwick B.P., Heath S.K., Williamson J., Obermayr F., Patel L., Sheer D., Fritsch A.M.;

RT "The human homologue of the ninjurin gene maps to the candidate region of hereditary sensory neuropathy type I (HSNI)."

RT Genomics 47:58-63(1998).

RN [3]

RP SEQUENCE FROM N.A.

RX TISSUE-Lung;

RA Strausberg R.;

RT Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.

RN [4]

RP CHARACTERIZATION.

RX MEDLINE-97407927: PubMed-9261151;

RA Arai T., Zimonjic D.B., Popescu N.C., Milderand J.;

RT "Mechanism of homophilic binding mediated by ninjurin, a novel widely expressed adhesion molecule."

RT J. Biol. Chem. 272:21373-21380(1997).

CC -1- FUNCTION: HOMOPHILIC CELL ADHESION MOLECULE THAT PROMOTES AXONAL GROWTH. MAY PLAY A ROLE IN NERVE REGENERATION AND IN THE FORMATION AND FUNCTION OF OTHER TISSUES. CELL ADHESION REQUIRES DIVALENT CATIONS.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED IN BOTH ADULT AND EMBRYONIC TISSUES, PRIMARILY THOSE OF EPITHELIAL ORIGIN.

CC -1- INDUCTION: BY NERVE INJURY BOTH IN DORSAL ROOT GANGLION NEURONS AND IN SCHWANN CELLS.

CC -1- SIMILARITY: BELONGS TO THE NINJURIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

DR EMBL: U72661; AAB17560.1; -

DR EMBL: U91512; AAC14593.1; -

DR EMBL: AF029251; AAC39574.1; -

DR EMBL: BC004440; AA04440.1; -

DR Gene: HGNC:7824; NINJ1.

DR MIM: 602062; -

DR Cell adhesion; Transmembrane.

FT DOMAIN 1 80 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 81 101 POTENTIAL.

FT DOMAIN 102 120 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 121 141 POTENTIAL.

FT DOMAIN 142 152 EXTRACELLULAR (POTENTIAL).

FT CONFLICT 110 110 D -> A (IN REF. 3).

SO SEQUENCE 152 AA; 16389 MW; FEACAD8001597855 CRC64;

Query Match

Best Local Similarity 40.0%; Score 18; DB 1; Length 152;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

DB 44 KKSABESMD 53

OY 3 KXXXXXXD 12

DB 11 11

DB 44 KKSABESMD 53

RESULT 3

NIJL\_MOUSE STANDARD: PRT: 152 AA.

AC 070131:

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Ninjurin 1 (Nerve injury-induced protein 1).

GN NINJ1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

ON [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-98126436: PubMed-9465296;

RA Chadwick B.P., Heath S.K., Williamson J., Obermayr F., Patel L., Sheer D., Fritsch A.M.;

RT "The human homologue of the ninjurin gene maps to the candidate region of hereditary sensory neuropathy type I (HSNI)."

RT Genomics 47:58-63(1998).

RN [2]

RP SEQUENCE FROM N.A.

RX Moon A.R., Kim J.W., Hong Y.M., Oh G.T., Chang S.Y., Lee K.S., Choe I.S.;

RT "Mus musculus ninjurin genomic DNA."

RT Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.

CC -1- FUNCTION: HOMOPHILIC CELL ADHESION MOLECULE THAT PROMOTES AXONAL GROWTH. MAY PLAY A ROLE IN NERVE REGENERATION AND IN THE FORMATION AND FUNCTION OF OTHER TISSUES. CELL ADHESION REQUIRES DIVALENT CATIONS (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- INDUCTION: BY NERVE INJURY.

CC -1- SIMILARITY: BELONGS TO THE NINJURIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

DR EMBL: U91513; AAC14594.1; -

DR EMBL: AF219626; AAC32161.1; -

DR MGD: MGI:1196617; NINJ1.

DR Cell adhesion; Transmembrane.

FT DOMAIN 1 79 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 80 100 POTENTIAL.

FT DOMAIN 101 120 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 121 141 POTENTIAL.

FT DOMAIN 142 152 EXTRACELLULAR (POTENTIAL).

SO SEQUENCE 152 AA; 16555 MW; E261CB447BC0A2E6 CRC64;

Query Match

Best Local Similarity 40.0%; Score 18; DB 1; Length 152;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KXXXXXXD 12

DB 11 11

DB 44 KKSAAESMLD 53

RESULT 4

ID NINJ1\_RAT STANDARD; PRT: 152 AA.

AC P70617;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE NINjurin 1 (Nerve injury-induced protein 1).

GN NINJ1.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

ON NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96374367; Pubmed=8780658;

RA Arai T., Milbrandt J.;

RT "Ninjurin, a novel adhesion molecule, is induced by nerve injury and promotes axonal growth."

RL Neuron 17:353-361(1996).

CC -1- FUNCTION: HOMOPHILIC CELL ADHESION MOLECULE THAT PROMOTES AXONAL GROWTH. MAY PLAY A ROLE IN NERVE REGENERATION AND IN THE FORMATION AND FUNCTION OF OTHER TISSUES. CELL ADHESION REQUIRES DIVALENT CATIONS (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- SIMILARITY: BELONGS TO THE NINJURIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: U72660; AAB17559.1; -

KW Cell adhesion; Transmembrane.

FT DOMAIN 1 79 EXTRACELLULAR (POTENTIAL).

FT TRANSSEM 80 100 POTENTIAL.

FT DOMAIN 101 120 CYTOPLASMIC (POTENTIAL).

FT TRANSSEM 121 141 POTENTIAL.

FT DOMAIN 142 152 EXTRACELLULAR (POTENTIAL).

SO SEQUENCE 152 AA; 16539 MW; 414B3B9DE1807E80 CRC64;

Query Match 64.3%; Score 18; DB 1; Length 152;

Best Local Similarity 40.0%; Pred. No. 6.5;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KXXXXXXLD 12

DB 44 KKSAAESMLD 53

RESULT 5

ID YML3\_PART2 STANDARD; PRT: 169 AA.

AC P15614;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 01-APR-1990 (Rel. 14, Last annotation update)

DE Hypothetical 20.2 kDa protein (ORF13).

OS Paramacium tetraurelia.

OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Penicillidae; Paramacium.

ON NCBI\_TaxID=5888;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Stock 51;

RX MEDLINE=90174913; Pubmed=2308823;

RA Pritchard A.E., Sellhammer J.J., Mahalingam R., Sable C.L., Venu S.E., Cummings D.J.;

RT "Nucleotide sequence of the mitochondrial genome of Paramacium.";

RL Nucleic Acids Res. 18:173-180(1990).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: X15917; CA34057.1; -

DR PIR: S07748; S07748.

KW Hypothetical protein; Mitochondrion.

SO SEQUENCE 169 AA; 20303 MW; D991AD99A9EBA174 CRC64;

Query Match 64.3%; Score 18; DB 1; Length 169;

Best Local Similarity 40.0%; Pred. No. 7.2;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KXXXXXXLD 12

DB 36 KKTSTYSKLD 45

RESULT 6

ID YCDC\_ECOLI STANDARD; PRT: 212 AA.

AC P75899;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical transcriptional regulator ycdc.

GN YCDC OR B1013.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

ON NCBI\_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=K12 / MG1655;

RY MEDLINE=97426617; Pubmed=9278503;

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Mau B., Shao Y.;

RA "The complete genome sequence of Escherichia coli K-12.";

RT Science 277:1453-1474(1997).

RL [2]

RP SEQUENCE FROM N.A.

RC STRAIN=K12;

RX MEDLINE=97061202; Pubmed=8905232;

RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A., Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K., Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K., Mori H., Motomura K., Nakamura Y., Nishio Y., Saito N., Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y., Yano M., Horikuchi T.;

RA "A 718-kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 12.7-28.0 min region on the linkage map.";

RT DNA Res. 3:137-155(1996).

RL -1- SIMILARITY: BELONGS TO THE YETR/ACRR FAMILY OF TRANSCRIPTIONAL REGULATORS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC or send an email to [license@sls-sib.ch](mailto:license@sls-sib.ch)).

CC -----

DR EMBL: AE000203; AAC74098.1; -

DR EMBL: D90738; BAA35790.1; -

DR EMBL: EG12301; YCCG.

DR InterPro: IPR001647; HTH\_Tetr.

DR Pfam: PF00440; tetr.1.

DR PROSITE: PS01081; HTH\_TETR\_FAMILY; FALSE\_NEG.

KW Hypothetical protein; Transcription regulation; DNA-binding;

KM Complete proteome.

FT DNA\_BIND

FT 39 58 H-T-H MOTIF (POTENTIAL).

SQ SEQUENCE 212 AA; 23687 MW; E2459B85DFAC277A CRC64;

QY 3 KKKXXXXXD 12

DB 20 KKAISALD 29

RESULT 7

TFOX\_HAEIN

AC P43779; STANDARD; PRT; 217 AA.

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE DNA transformation protein tfox (competence activator) (Protein sxy).

GN TFOX OR SXY OR HT0601.

OS Haemophilus influenzae.

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

OC Haemophilus.

OX NCBI\_TaxID=727;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-Rd.

RX MEDLINE-95241551; PubMed-7724607;

RA Zully J.J., Barck G.J.;

RT "Identification of a DNA transformation gene required for com101A+ expression and supertransformer phenotype in Haemophilus influenzae."

RT Proc. Natl. Acad. Sci. U.S.A. 92:3616-3620(1995).

RL [2]

RN SEQUENCE FROM N.A.

RP STRAIN-Rd / KW20 / ATCC 51907;

RC MEDLINE-95350630; PubMed-7542800;

RX Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F., Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M., McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D., Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M., Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D., Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C., Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghegan N.S.M., Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O., Venter J.C.;

RA "Whole-genome random sequencing and assembly of Haemophilus influenzae Rd."

RT Science 269:496-512(1995).

RL [3]

RN IDENTIFICATION OF TFOX AS SXY AND VARIANT SXY-1.

RC MEDLINE-95050244; PubMed-7961436;

RX Williams P.M., Bannister L.A., Redfield R.J.;

RA "The Haemophilus influenzae sxy-1 mutation is in a newly identified gene essential for competence."

RT J. Bacteriol. 176:6789-6794(1994).

RL -1- FUNCTION: REQUIRED FOR DNA TRANSFORMATION (LATE COMPETENCE-SPECIFIC GENES).

CC -1- SIMILARITY: TO E.COLI YCCR.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>).

CC or send an email to [license@sls-sib.ch](mailto:license@sls-sib.ch)).

CC -----

DR EMBL: U13205; AAC43320.1; -

DR EMBL: U13205; AAC43321.1; -

DR EMBL: U32741; AAC22258.1; -

DR TIGR: HT0601; -

KW Competence; Transcription regulation; Activator;

KM Alternative initiation; Complete proteome.

FT CHAIN 1 217

FT 1 217 DNA TRANSFORMATION PROTEIN TFOX-SHORT.

FT INIT MET 40 217 FOR TFOX-SHORT.

FT VARIANT 19 19 V -> I (IN SXY-1: CAUSES A 100 X TO 1000 X INCREASE IN SPONTANEOUS NATURAL COMPETENCE).

FT W -> L (IN REF. 1).

FT CONFLICT 203 203

SQ SEQUENCE 217 AA; 25074 MW; 4A6B30E64008CC3F CRC64;

QY 3 KKKXXXXXD 12

DB 161 KKGSGATMD 170

RESULT 8

ORN\_HUMAN

AC Q9Y3B8; Q9UEY7; STANDARD; PRT; 237 AA.

DT 30-MAY-2000 (Rel. 39, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Oligoribonuclease, mitochondrial precursor (EC 3.1.1.1) (Small fragment nuclease) (CGI-114).

GN SFN.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (SHORT ISOFORM).

RC MEDLINE-20272150; PubMed-10810993;

RX Lai C.-H., Chou C.-Y., Chang L.-Y., Liu C.-S., Lin W.-C.;

RA "Identification of novel human genes evolutionarily conserved in Caenorhabditis elegans by comparative proteomics."

RT genome Res. 10:703-713(2000).

RL [2]

RN SEQUENCE FROM N.A. (LONG ISOFORM).

RP TISSUE-Kidney.

RC MEDLINE-21154917; PubMed-11230166;

RX Wiemann S., Well B., Wellenreuther R., Gassenhuber J., Glassl S., Ansorge W., Boecker M., Bloeker H., Bauersachs S., Blum H., Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N., Lewes H.-W., Ottenweider B., Oberwalder B., Tanpe J., Heubner D., Wambutt R., Korn B., Klein M., Poustka A.;

RA "Towards a catalog of human genes and proteins: sequencing and analysis of 500 novel complete protein coding human cDNAs."

RT genome Res. 11:422-435(2001).

RL [3]

RN CHARACTERIZATION.

RC MEDLINE-20408924; PubMed-10851236;

RX Nguyen L.H., Erdberger J.P., Root J., Wilson D.M. III;

RA "The human homolog of Escherichia coli Orn degrades small single-stranded RNA and DNA oligomers."

RT J. Biol. Chem. 275:25900-25906(2000).

RL -1- FUNCTION: 3'-5' EXORIBONUCLEASE SPECIFIC FOR SMALL OLIGORIBONUCLEOTIDES. CC ACTIVE ON SMALL (PRIMARILY </=5



CC NUCLEOTIDES IN LENGTH) SINGLE-STRANDED RNA AND DNA OLIGOMERS. MAY  
 CC HAVE A ROLE FOR CELLULAR NUCLEOTIDE RECYCLING.  
 CC -1- COFACTOR: MANGANESE.  
 CC -1- SUBUNIT: HOMOTETRAMER (PROBABLE).  
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL (LONG ISOFORM) AND NUCLEAR  
 CC (SHORT ISOFORM).  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM/SEN-ALPHA (SHOWN  
 CC HERE) AND A SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: BELONGS TO THE OLIGORIBONUCLEASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>  
 CC or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch)).  
 CC -----  
 CC EMBL; AF151872; A034109.1; -  
 DR EMBL; AL110239; CAB53690.1; -  
 DR InterPro: IPR000520; Exonuclease.  
 DR Pfam: PF00929; Exonuclease; 1.  
 DR SMART: SM00479; EXOIII.1.  
 KM Hydroxylase; Exonuclease; Nuclease; Mitochondrion; Transit peptide;  
 KW Nuclear protein; Manganese; Alternative splicing.  
 FT TRANSIT 1  
 FT CHAIN ? 237  
 FT ACT\_SITE 164 164 POTENTIAL.  
 FT VARSPIC 1 38 MISSING (IN ISOFORM 2).  
 FT MTAGN 168 168 D->A: ABOLISHES ACTIVITY.  
 FT CONFLICT 103 103 R->K (IN REF. 1).  
 SQ SEQUENCE 237 AA; 26861 MW; BF093A8907BC13BF CRC64;  
 QY 3 KKKXXXXXLD 12  
 Db 189 KKAASHRALD 198  
 RESULT 9  
 GPD\_CHLTR STANDARD; PRT; 247 AA.  
 ID GPD\_CHLTR STANDARD; PRT; 247 AA.  
 AC P10560; P08787; Q46433;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-JUN-1989 (Rel. 11, Last sequence update)  
 DT 15-OCT-2001 (Rel. 40, Last annotation update)  
 DE Virulence plasmid protein pGP6-D (Protein P-10).  
 OS Chlamydia trachomatis.  
 OC Plasmid pLG440; Plasmid pCHL1; and Plasmid pCTT1.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_Taxid-813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-L2/434/Bu; PLASMID-pLG440;  
 RX MEDLINE-89013895; PubMed-2845228;  
 RA Comanducci M., Ricci S., Ratti G.;  
 "The structure of a plasmid of Chlamydia trachomatis believed to be  
 RT required for growth within mammalian cells.";  
 RL Mol. Microbiol. 2:531-538(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-L1/440/LN; PLASMID-pLG440;  
 RX MEDLINE-8823398; PubMed-2836808;  
 RA Hatt C., Ward M.E., Clarke I.N.;  
 "Analysis of the entire nucleotide sequence of the cryptic plasmid of  
 RT Chlamydia trachomatis serovar L1. Evidence for involvement in DNA  
 RL Nucleic Acids Res. 16:4053-4067(1988).  
 RN [3]  
 RP SEQUENCE FROM N.A.

RC STRAIN-G0/86 / Serotype D; PLASMID-pCHL1;  
 RX MEDLINE-90301796; PubMed-2194229;  
 RA Comanducci M., Ricci S., Cevenini R., Ratti G.;  
 RT "Diversity of the Chlamydia trachomatis common plasmid in biovars  
 RL with different pathogenicity.";  
 RL Plasmid 23:149-154(1990).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Serotype B; PLASMID-pCTT1;  
 RX MEDLINE-8817106; PubMed-3444859;  
 RA Sriprakash K.S., Macavoy E.S.;  
 RT "Characterization and sequence of a plasmid from the trachoma biovar  
 RL of Chlamydia trachomatis.";  
 RL Plasmid 18:205-214(1987).  
 CC -1- MISCELLANEOUS: PG6-D IS REQUIRED FOR GROWTH WITHIN MAMMALIAN  
 CC CELLS.  
 CC -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF PLASMID pLG440.  
 CC -1- SIMILARITY: BELONGS TO THE UPF0137 (PG6-D) FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>  
 CC or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch)).  
 CC -----  
 CC EMBL; X07547; CA30426.1; -  
 DR EMBL; X06707; CA29897.1; -  
 DR EMBL; J03321; AA91574.1; -  
 DR EMBL; M19487; AAB02591.1; -  
 DR PIR; S01924; S01924.  
 DR PIR; S00794; S00794.  
 DR InterPro: IPR005350; UPF0137.  
 DR Pfam; PF03677; UPF0137; 1.  
 KM Plasmid.  
 FT VARIANT 9  
 FT VARIANT 47 47 N->D (IN PLASMIDS PCHL1 AND PCTT1).  
 FT VARIANT 183 186 Q->R (IN PLASMIDS PCHL1 AND PCTT1).  
 FT VARIANT 194 194 VLQD->CWTE (IN PLASMID PCTT1).  
 FT VARIANT 205 205 K->T (IN PLASMID PCTT1).  
 FT VARIANT 205 205 S->L (IN PLASMID PCTT1).  
 FT VARIANT 222 222 M->I (IN PLASMID PCTT1).  
 SQ SEQUENCE 247 AA; 28307 MW; F1FA11F29037C221 CRC64;  
 QY 3 KKKXXXXXLD 12  
 Db 13 KKNQTAASLD 22  
 RESULT 10  
 VPQ\_BPP2 STANDARD; PRT; 344 AA.  
 ID VPQ\_BPP2 STANDARD; PRT; 344 AA.  
 AC P25480;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Presumed portal vertex protein (GPO).  
 GN O.  
 OS Bacteriophage P2.  
 CC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;  
 CC P2-like viruses.  
 OX NCBI\_Taxid-10679;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE-92115571; PubMed-1837355;  
 RX Linderoth N.A., Ziermann R., Hagggaard-Ljungquist E., Christie G.E.,  
 RA Calendar R.;  
 RT "Nucleotide sequence of the DNA packaging and capsid synthesis genes  
 of bacteriophage P2.";

CC Nucleic Acids Res. 19:7207-7214(1991).  
 CC -1- FUNCTION: THE Q, P AND M PROTEINS ARE NEEDED TO PACKAGE DNA INTO  
 CC PROHEADS AND FOR THE CONVERSION OF PROHEADS TO CAPSID.  
 CC -1- SIMILARITY: STRONG, TO PHAGE HP1 PROTEIN ORF15 AND TO THE E. COLI  
 CC HYPOTHETICAL 36.8 kDa PROTEIN (ORF5) IN RETRON EC67.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF063097; AAD03268.1; -  
 DR PIR; S16410; S16410.  
 DR PIR; S22796; S22796.  
 KM Capsid assembly.  
 SQ SEQUENCE 344 AA; 39114 MW; FAD61987530029A7 CRC64;  
 Query Match 64.3%; Score 18; DB 1; Length 344;  
 Best Local Similarity 40.0%; Pred. No. 14;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 OY 3 KKKXXXXXD 12  
 DB 270 KKASADLD 279  
 RESULT 11  
 ID FLIC1\_PROM1 STANDARD; PRT; 365 AA.  
 AC P42272;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Flagellin 1.  
 GN FLIC1.  
 OS Proteus mirabilis.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Proteus.  
 OX NCBI\_TaxID=584;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-29.  
 RC STRAIN-BB2000;  
 RX MEDLINE-95011656; PubMed-7926835;  
 RA Belas R., Flaherty D.;  
 RT "Sequence and genetic analysis of multiple flagellin-encoding genes  
 RT from Proteus mirabilis.";  
 RL Gene 148:33-41(1994).  
 CC -1- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO  
 CC FORM THE FILAMENTS OF BACTERIAL FLAGELLA.  
 CC -1- INDUCTION: ALTHOUGH SWIMMER CELLS HAVE ONLY A FEW FLAGELLA, THE  
 CC ELONGATED SWARMER CELLS ARE PROPOSEDLY COVERED BY THOUSANDS OF NEW  
 CC FLAGELLA SYNTHESIZED SPECIFICALLY IN RESPONSE TO GROWTH ON  
 CC SURFACES OR IN HIGHLY VISCOUS LIQUIDS.  
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF221596; AAA62396.1; -  
 DR InterPro; IPR001492; Flagellin.  
 DR InterPro; IPR001029; Flagellin\_C.  
 DR Pfam; PF00669; Flagellin\_N; 1.  
 DR Pfam; PF00700; Flagellin\_C; 1.  
 DR PRINTS; PR00207; FLAGELLIN.  
 DR ProDom; PD000316; Flagellin\_C; 1.

KW Flagella; Multigene family.  
 SQ SEQUENCE 365 AA; 39093 MW; 2A23106F040DB522 CRC64;  
 Query Match 64.3%; Score 18; DB 1; Length 365;  
 Best Local Similarity 40.0%; Pred. No. 15;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 OY 3 KKKXXXXXD 12  
 DB 197 KKNMAATLD 206  
 RESULT 12  
 ID Y349\_MYCPN STANDARD; PRT; 413 AA.  
 AC P75253;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein MG349 homolog (G12\_orf413).  
 GN MPN525 OR MP317.  
 OS Mycoplasma pneumoniae.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2104;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 29342 / M129;  
 RX MEDLINE-97105885; PubMed-8948633;  
 RA Himmelfreith R., Hilbert H., Pliagens H., Pirkl E., Li B.-C.,  
 RA Herrmann R.;  
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
 RT pneumoniae.";  
 RL Nucleic Acids Res. 24:4420-4449(1996).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE000029; AAB95965.1; -  
 DR Hypothetical protein; Complete proteome.  
 KM SEQUENCE 413 AA; 48246 MW; A4CP2C9802316F23 CRC64;  
 Query Match 64.3%; Score 18; DB 1; Length 413;  
 Best Local Similarity 40.0%; Pred. No. 17;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 OY 3 KKKXXXXXD 12  
 DB 10 KKATSPSGID 19  
 RESULT 13  
 ID ARP3\_SCHPO STANDARD; PRT; 427 AA.  
 AC P32390;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Actin-like protein 3.  
 GN ACT2 OR SPAC630.03.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-92108083; PubMed-1729722;  
 RA Lees-Miller J.P., Henry G., Helfman D.M.;

\* Identification of act2, an essential gene in the fission yeast Schizosaccharomyces pombe that encodes a protein related to actin. [2]  
Proc. Natl. Acad. Sci. U.S.A. 89:80-83(1992).

RP SEQUENCE FROM N.A.  
RC STRAIN-972;  
RX MEDLINE-21648401; PubMed-11859360;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A., Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S., Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M., Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A., Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G., Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K., James K., Jones L., Jones M., Leather S., McDonald S., Mclean J., Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C., Oliver K., O'Neill S., Pearson D., Quail M.A., Rabblowitsch E., Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S., Skelton J., Simmonds M., Squares R., Squares S., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Volkart G., Aert R., Robben J., Grymonprez B., Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H., Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M., Eger P., Zimmermann W., Wedler H., Wandutt R., Purnelle B., Corfeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Kottler S., Galbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M., Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G., Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J., Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L., Cerritelli L., Lowe T., McCombie W.R., Paulsen I., Potashkin J., Siparowski G.V., Ussery D., Barrett B.G., Nurse P.;  
RT The genome sequence of Schizosaccharomyces pombe.\*;  
RL Nature 415:871-880(2002).

-1- FUNCTION: PART OF A COMPLEX IMPLICATED IN THE CONTROL OF ACTIN POLYMERIZATION IN CELLS (BY SIMILARITY). MAY BE INVOLVED IN CYTOKINESIS.  
CC -1- SUBUNIT: BELONGS TO A COMPLEX COMPOSED OF ARP2, ARP3, P41-ARC, P34-ARC, P21-ARC, P20-ARC AND P16-ARC (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE ACTIN FAMILY. ARP3 SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC  
CC  
CC EMBL; M81068; -; NOT\_ANNOTATED\_CDS.  
DR EMBL; AL109832; CAB52725.1; -;  
DR PIR; A41790; A41790.  
DR InterPro: IPR004000; Actin\_like.  
DR Pfam: PF000022; actin; 1.  
DR SMART; SM00268; ACTIN; 1.  
DR PROSITE; PS01132; ACTINS\_ACT\_LIKE; 1.  
DR Structural Protein; Cytoskeleton.  
KW SEQUENCE 427 AA; 47373 MW; C7909FEE544789B CRC64;  
SQ

Query Match 64.3%; Score 18; DB 1; Length 427;  
Best Local Similarity 40.0%; Pred. NO. 17;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KKKKXXIXD 12  
Db 82 KKASAGYSD 91

RESULT 14  
XYNA\_STRLI STANDARD; PRT; 477 AA.  
AC P26514; P96464;  
DT 01-APR-1992 (Rel. 23, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Endo-1,4-beta-xylanase A precursor (EC 3.2.1.8) (Xylanase A)  
GN (1,4-beta-D-xylan xylanohydrolase A).  
OS Streptomyces lividans.  
OC Bacteria; Actinobacteria; Actinobacteriales (class); Actinobacteridae;  
OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID:1916;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 42-92.  
RC STRAIN-66 / 1326;  
RX MEDLINE-92077439; PubMed-1743521;  
RA Shareck F., Roy C., Yaguchi M., Morosoli R., Kluepfel D.;  
RT "Sequences of three genes specifying xylanases in Streptomyces lividans.";  
RL Gene 107:75-82(1991).  
RN [2]  
RP REVISIONS TO 20 AND 140-141.  
RC STRAIN-66 / 1326;  
RA Shareck F.;  
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 42-340.  
RX MEDLINE-94342232; PubMed-8063693;  
RA Derwenda U., Swenson L., Green R., Wei Y.Y., Morosoli R., Shareck F., Kluepfel D., Derwenda Z.S.;  
RT Crystal structure, at 2.6-A resolution, of the Streptomyces lividans xylanase A, a member of the F family of beta-1,4-D-glucanases.";  
RL J. Biol. Chem. 269:20811-20814(1994).

-1- FUNCTION: CONTRIBUTES TO HYDROLYSE Hemicellulose, THE MAJOR COMPONENT OF PLANT CELL-WALLS. XLNA AND XLNB SEEM TO ACT SEQUENTIALLY ON THE SUBSTRATE TO YIELD xylobiose AND xylose AS CARBON SOURCES.  
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-xylosidic linkages in xylans.  
CC -1- PATHWAY: Xylan degradation.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY F (FAMILY 10 OF GLYCOSYL HYDROLASES).  
CC -1- SIMILARITY: CONTAINS 1 RICIN B-TYPE LECTIN DOMAIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC  
CC  
CC EMBL; M64551; AAC26525.1; -;  
DR PIR; JS0589; JS0589.  
DR PDB; 1XAS; 31-MAY-95.  
DR InterPro: IPR001000; Glyco\_hydro\_10.  
DR InterPro: IPR000772; Ricin\_B\_lectin.  
DR Pfam: PF00331; Glyco\_hydro\_10; 1.  
DR Pfam: PF00652; Ricin\_B\_lectin; 3.  
DR PRINTS; PR00134; GLHYDRLASE10.  
DR SMART; SM00458; RICIN; 1.  
DR PROSITE; PS00591; GLYCOSYL\_HYDROL\_F10; 1.  
DR PROSITE; PS00231; RICIN\_B\_LLECTIN; 1.  
KW Xylan degradation; Hydrolyase; Glycosidase; Signal; Lectin; 3D-structure.  
FT SIGNAL 1 41  
FT CHAIN 42 477 ENDO-1,4-BETA-XYLANASE A.  
FT DOMAIN 361 477 RICIN B-TYPE LECTIN.  
FT ACT\_SITE 169 169 PROTON DONOR.  
FT ACT\_SITE 277 277 NUCLEOPHILE.  
SQ SEQUENCE 477 AA; 51162 MW; E14A7FE37BDC68CC CRC64;  
SQ

Query Match 64.3%; Score 18; DB 1; Length 477;  
Best Local Similarity 40.0%; Pred. NO. 19;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KXXXXXXXXD 12  
DB 330 KKAAYTAVLD 339

RESULT 15

YL34\_YEAST ID YL34\_YEAST STANDARD; PRT; 837 AA.

AC 007844;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE Hypothetical 93.1 kDa protein YL034C.

GN YL034C.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI\_Taxid=4932;

RP [1] SEQUENCE FROM N.A. Duesterhoeft A., Floeth M., Heuss-Neltzel D., Hilbert H., Moestl D.;

RL Submitted (MAY-1996) to the EMBL/Genbank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: Z73139; CAA97483.1; -.

DR SGD: S0003957; YL034C.

DR InterPro: IPR003593; AAA\_ATPase.

DR InterPro: IPR003599; AAA\_ATPase-centr.

DR InterPro: IPR003960; AAA\_sub.

DR Pfam: PF00004; AAA; 2.

DR SMART: SM00382; AAA; 1.

DR PROSITE: PS00674; AAA; 1.

KW Hypothetical protein: ATP-binding. Repeat.

FT NP\_BIND 246 253 ATP (POTENTIAL).

FT NP\_BIND 574 581 ATP (POTENTIAL).

FT NP\_BIND 93069 93069 MW: 13528890987103 CRC64;

SQ SEQUENCE 837 AA; 93069 MW: 13528890987103 CRC64;

Query Match 64.3%; Score 18; DB 1; Length 837;

Best Local Similarity 40.0%; Pred. NO. 33;

Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KXXXXXXXXD 12

DB 7 KKNSTSSLD 16

Search completed: June 5, 2003, 15:56:35

Job time : 8.81818 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: June 5, 2003, 15:32:58 ; Search time 34.9091 Seconds  
(without alignments)  
70.829 Million cell updates/sec

Title: US-09-150-947f-17  
Perfect score: 28  
Sequence: 1 XXXXXXXXXXXX 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

SPREMBL\_21.\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertedrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	67.9	284	3	09URX9
2	19	67.9	378	10	09ZVH8
3	19	67.9	378	10	08VXZ6
4	19	67.9	1365	12	037174
5	19	67.9	1365	12	065005
6	19	67.9	1365	12	08U248
7	18	64.3	46	2	09ZG37
8	18	64.3	108	2	08RRC9
9	18	64.3	153	2	09ANS7
10	18	64.3	160	4	08WUV5
11	18	64.3	161	16	09K132
12	18	64.3	171	10	004541
13	18	64.3	205	11	09D854
14	18	64.3	212	16	08X427
15	18	64.3	215	5	096609
16	18	64.3	276	8	P92729

17	18	64.3	278	16	08U709	08U709 agrobacteri
18	18	64.3	292	11	08R1E2	08R1E2 mus musculu
19	18	64.3	357	15	090F08	090F08 human immun
20	18	64.3	371	15	090K19	090K19 human immun
21	18	64.3	376	12	056872	056872 gallid hezp
22	18	64.3	385	15	0982J3	0982J3 human immun
23	18	64.3	543	5	044130	044130 caenorhabdl
24	18	64.3	663	8	09M069	09M069 physarum po
25	18	64.3	716	2	09REY3	09REY3 carobacter
26	18	64.3	716	2	046317	046317 carobacter
27	18	64.3	716	2	0930C9	0930C9 carobacter
28	18	64.3	765	4	0968N7	0968N7 homo sapien
29	18	64.3	765	4	0966Q7	0966Q7 homo sapien
30	18	64.3	796	4	096197	096197 homo sapien
31	18	64.3	841	2	044207	044207 agrobacteri
32	18	64.3	858	5	021853	021853 caenorhabdl
33	18	64.3	863	16	097LNO	097LNO clostridum
34	18	64.3	954	5	09G015	09G015 giardia lam
35	18	64.3	969	5	09U022	09U022 giardia lam
36	18	64.3	1032	5	019818	019818 caenorhabdl
37	18	64.3	1035	5	021380	021380 caenorhabdl
38	18	64.3	1117	5	08T1U6	08T1U6 dictyostell
39	18	64.3	1271	2	0935V9	0935V9 chlorobium
40	17	60.7	72	5	096726	096726 lutomyia l
41	17	60.7	109	17	028686	028686 archaeglob
42	17	60.7	119	17	0975E7	0975E7 sulfolobus
43	17	60.7	137	16	09PKF9	09PKF9 chlamydia m
44	17	60.7	137	16	084238	084238 chlamydia t
45	17	60.7	145	5	08SWF8	08SWF8 encephalit

## ALIGNMENTS

RESULT 1	PRELIMINARY:	PRT:	284 AA.
ID 09URX9	01-MAY-2000 (TREMBLrel. 13, Created)		
AC 09URX9	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE Hypothetical 31.9 kDa protein C890.05 in chromosome I.			
GN SPAC890.05.			
OS Schizosaccharomyces pombe (Fission yeast).			
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;			
OC Schizosaccharomycetales; Schizosaccharomycetaceae;			
OC Schizosaccharomyces.			
OX NCBI_Taxid=4896;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=972;			
RA Saunders D., Harris D., McDougall R.C., Rajandream M.A., Barrell B.G.;			
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.			
DR EMBL; AL133498; CAB63496.1; -			
DR InterPro: IPR000467; G_patch.			
DR Pfam: PF01585; G_patch; 1.			
DR SMART: SM00443; G_patch; 1.			
KW Hypothetical protein.			
FT DOMAIN 197..201			POLY-LYS.
FT DOMAIN 225..233			POLY-LYS.
SQ SEQUENCE 284 AA; A10P14DDE2FA0235 CRC64;			
Query Match	67.9%;	Score 19;	DB 3; Length 284;
Best Local Similarity	40.0%;	Pred. No. 20;	
Matches 4;	Conservative 0;	Mismatches 6;	Indels 0; Gaps 0;

QY	3	XXXXXXXXXD 12	
DB	179	KRTSGSALD 188	
RESULT 2			
Q9ZVH8			

```

ID 09ZYH8      PRELIMINARY;      PRT;      378 AA.
AC 09ZYH8:
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Putative elongation factor.
GN ATG38560
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Bentto M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum A.V.,
RA Buell C.R., Ketchum K.A., Lee J.U., Konning C.M., Koo H., Motil K.S.,
RA Cronin L.A., Shen M., Vanaken S.E., Umeyam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhagen G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana";
RL Nature 402:761-768 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AC005499; AAC67362.1; -.
DR HSP; F23193; 1FF1.
DR InterPro; IPR001222; TFIIS.
DR InterPro; IPR003618; TFS2_centre.
DR InterPro; IPR003617; TFS2_N.
DR Pfam; PF01096; TFIIS; 1.
DR SMART; SM00510; TFS2M; 1.
DR SMART; SM00509; TFS2N; 1.
DR SMART; SM00440; znf_C2C2; 1.
DR PROSITE; PS00466; TFIIS; 1.
KW Elongation factor.
SQ SEQUENCE 378 AA; 41979 MW; 9A176DC31FA8E885 CRC64;

Query Match      67.9%; Score 19; DB 10; Length 378;
Best Local Similarity 40.0%; Pred. No. 26;
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KKKXXXXXLD 12
DB 13 KKAADAAALD 22

RESULT 3
OBVXZ6
AC 08VXZ6:      PRELIMINARY;      PRT;      378 AA.
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative elongation factor.
GN AT2638560
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,

```

```

RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Natusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinozaki A., Shinozaki K.,
RA Davis R.W., Ecker J.R., Theologis A.;
RT "Full length cDNA of gene At2g38560 (GI:15224901)";
RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF074322; AAL67018.1; -.
DR InterPro; IPR001222; TFIIS.
DR InterPro; IPR003618; TFS2_centre.
DR InterPro; IPR003617; TFS2_N.
DR Pfam; PF01096; TFIIS; 1.
DR SMART; SM00510; TFS2M; 1.
DR SMART; SM00509; TFS2N; 1.
DR SMART; SM00440; znf_C2C2; 1.
DR PROSITE; PS00466; TFIIS; UNKNOWN_1.
KW Elongation factor.
SQ SEQUENCE 378 AA; 42011 MW; 045EE924FBDCCB CRC64;

Query Match      67.9%; Score 19; DB 10; Length 378;
Best Local Similarity 40.0%; Pred. No. 26;
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KKKXXXXXLD 12
DB 13 KKAADAAALD 22

RESULT 4
O37174
AC 037174:      PRELIMINARY;      PRT;      1365 AA.
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Replicase.
OS Bamboo mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potexvirus.
OX NCBI_TaxID=35286;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S;
RA Liao J.T., Hsu Y.H.;
RT "Full-length sequence of infectious bamboo mosaic virus genomic RNA.";
RL Submitted (AUG-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF018156; AAB70562.1; -.
DR InterPro; IPR001788; RNA_dep_RNAPol2.
DR InterPro; IPR006066; Viral_helicase1.
DR Pfam; PF00978; RNA_dep_RNAPol2; 1.
DR Pfam; PF01443; Viral_helicase1; 1.
SQ SEQUENCE 1365 AA; 155219 MW; CA54D8F5997F158A CRC64;

Query Match      67.9%; Score 19; DB 12; Length 1365;
Best Local Similarity 40.0%; Pred. No. 79;
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KKKXXXXXLD 12
DB 1302 KKTATATVALD 1311

RESULT 5
O65005
AC 065005:      PRELIMINARY;      PRT;      1365 AA.
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ORF1-155k.
OS Bamboo mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potexvirus.
OX NCBI_TaxID=35286;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN-BAMV-O;  
 RA Hsu Y.-H.;  
 RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BAMV-O;  
 RX MEDLINE-94358760; PubMed-8077956;  
 RA Lin N., Lin B., Lo N., Hu C., Chow T., Hsu Y.;  
 RT "Nucleotide sequence of the genomic RNA of bamboo mosaic potexvirus.";  
 RL J. Gen. Virol. 75:2513-2518(1994).  
 DR EMBL: D26017; BAA05033.1; -;  
 DR Interpro: IPR001788; RNA\_dep\_RNAPol2.  
 DR Interpro: IPR000606; Viral\_helicase1.  
 DR Pfam: PF01443; Viral\_helicase1; 1.  
 DR SEQUENCE 1365 AA; 155129 MW; 35449D5E80026633 CRC64;  
 SQ  
 Query Match 67.98; Score 19; DB 12; Length 1365;  
 Best Local Similarity 40.08; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 3 KKKXXXXXLD 12  
 Db 1302 KKTATAYALD 1311  
 RESULT 6  
 ID 08U248 PRELIMINARY; PRT; 1365 AA.  
 AC 08U248;  
 DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Putative, putative, putative, putative, and putative genes.  
 OS Bamboo mosaic virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potexvirus.  
 OX NCBI\_TaxID=35286;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BAMV-V;  
 RA Yang C.-C., Liu J.-S., Lin C.-P., Lin N.-S.;  
 RT "Nucleotide sequence and phylogenetic analysis of a bamboo mosaic  
 RT potexvirus isolate from common bamboo (Bambusa vulgaris McClure).";  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: L77963; AAL40030.1; -;  
 DR Interpro: IPR001788; RNA\_dep\_RNAPol2.  
 DR Interpro: IPR000606; Viral\_helicase1.  
 DR Pfam: PF00978; RNA\_dep\_RNAPol2; 1.  
 DR Pfam: PF01443; Viral\_helicase1; 1.  
 DR SEQUENCE 1365 AA; 155222 MW; B542C00E2ECFC990 CRC64;  
 SQ  
 Query Match 67.98; Score 19; DB 12; Length 1365;  
 Best Local Similarity 40.08; Pred. No. 79;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 3 KKKXXXXXLD 12  
 Db 1302 KKTATAYALD 1311  
 RESULT 7  
 ID 09ZG37 PRELIMINARY; PRT; 46 AA.  
 AC 09ZG37;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Hypoetical 5.2 kDa protein (Fragment).  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiales; Chlamydiales; Chlamydia.  
 OX NCBI\_TaxID=813;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN-L2 434B;  
 RA Wang L., Steenburg S.D., Zheng Y., Larsen S.H.;  
 RT "Gene identification of Chlamydia trachomatis by random DNA  
 RT sequencing.";  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF087330; AAD04105.1; -;  
 DR Interpro: IPR005350; UPF0137.  
 DR Pfam: PF03677; UPF0137; 1.  
 DR Hypoetical protein; Plasmid.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 46 AA; 5214 MW; D2FDAC41C1960867 CRC64;  
 Query Match 64.38; Score 18; DB 2; Length 46;  
 Best Local Similarity 40.08; Pred. No. 12;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 3 KKKXXXXXLD 12  
 Db 21 KKNQTAASLD 30  
 RESULT 8  
 ID 08RTC9 PRELIMINARY; PRT; 108 AA.  
 AC 08RTC9;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Putative portal vertex protein GPO (Fragment).  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CH014;  
 RA Pradel N., Leroy-Settin S., Livrelli V.;  
 RT "Sequences from Shiga toxin-producing Escherichia coli O91:H21.";  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF467512; AAL78348.1; -;  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 108 AA; 12028 MW; 350B505A7A03B5B5 CRC64;  
 Query Match 64.38; Score 18; DB 2; Length 108;  
 Best Local Similarity 40.08; Pred. No. 25;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 3 KKKXXXXXLD 12  
 Db 54 KKASADLDD 63  
 RESULT 9  
 ID 09ANS7 PRELIMINARY; PRT; 153 AA.  
 AC 09ANS7;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE LuxT.  
 OS Vibrio harveyi.  
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=669;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lin Y.H., Miyamoto C.M., Melgosh E.A.;  
 RT "Cloning and functional studies of a luxO regulator luxT from Vibrio  
 RT Blochum. Biophys. Acta 0:0-0(2001).  
 DR EMBL: AF321123; AAK09362.1; -;

SO SEQUENCE 153 AA; 17475 MW; 33837BBI2A2A1819B CRC64;  
 Query Match 64.3%; Score 18; DB 2; Length 153;  
 Best Local Similarity 40.0%; Pred. No. 34;  
 Matches 4; Conservative 0; Mismatches -6; Indels 0; Gaps 0;

OY 3 KXXXXXXXXX 12  
 DB 54 KKTDFTAALD 63

RESULT 10  
 Q8WV5 PRELIMINARY; PRT; 160 AA.  
 AC Q8WV5;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Hypothetical 17.0 kDa protein (fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=SKIN;  
 RA Strusberg R.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC019336; AAH19336.1;  
 KW Hypothetical protein.  
 FT NON\_TER  
 SQ SEQUENCE 160 AA; 17041 MW; 9F7B497AAC065DB8 CRC64;

Query Match 64.3%; Score 18; DB 4; Length 160;  
 Best Local Similarity 40.0%; Pred. No. 35;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KXXXXXXXXX 12  
 DB 52 KKSAAESMID 61

RESULT 11  
 Q9KL32 PRELIMINARY; PRT; 161 AA.  
 AC Q9KL32;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Transcriptional regulator, TetR family.  
 GN VCA0917.  
 OS Vibrio cholerae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=666;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=EL TOR N16961 / SEROTYPE O1;  
 RX MEDLINE=20406833; PubMed=10952301;  
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,  
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragol I., Sellers P.,  
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
 RA Fraser C.M.;  
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
 cholerae";  
 RL Nature 406:477-483(2000).  
 DR EMBL; AE004419; AAF96814.1;  
 DR TIGR; AC009117;  
 KW Complete proteome.  
 SQ SEQUENCE 161 AA; 18749 MW; AA78003D189B4288 CRC64;

Query Match 64.3%; Score 18; DB 16; Length 161;  
 Best Local Similarity 40.0%; Pred. No. 35;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KXXXXXXXXX 12  
 DB 62 KKTDFASALD 71

RESULT 12  
 O04541 PRELIMINARY; PRT; 171 AA.  
 AC O04541;  
 DT 01-JUL-1997 (TREMBLrel. 04, Created)  
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE F20P5.23 protein.  
 GN F20P5.23.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC Eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Osborne B.I., Vysotskaya V.S., Tortum M., Yu G., Oji O., Buehler E.,  
 RA Conway A.B., Conway A.R., Dewar K., Feng J., Kim C., Kuritz D., Li Y.,  
 RA Shin F., Sun H., Davis R.W., Ecker J.R., Federspiel N.A.,  
 RA Theologis A.;  
 RT "The sequence of BAC F20P5 from Arabidopsis thaliana chromosome 1";  
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Theologis A.;  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC002062; AAB61109.1;  
 DR InterPro; IPR003822; PAH.  
 DR Pfam; PF02671; PAH; 2.  
 SQ SEQUENCE 171 AA; 19992 MW; 691285E59734CA73 CRC64;

Query Match 64.3%; Score 18; DB 10; Length 171;  
 Best Local Similarity 40.0%; Pred. No. 37;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KXXXXXXXXX 12  
 DB 16 KKLTTTDAALD 25

RESULT 13  
 Q9D854 PRELIMINARY; PRT; 205 AA.  
 AC Q9D854;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE 1810038D15RIK protein (RIKEN CDNA 1810038D15 gene).  
 GN 1810038D15RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=PANCREAS;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,  
 RA Saito T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,



RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kiehl P., Lewis S., Matsuo Y., Nakado I., Pesole G., Quackenbush J.,  
 RA Schiraldi L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barag G.,  
 RA Blake J., Boffelli D., Bojunga N., Carlini P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gunglrich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Yushanaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kontsuki S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.,  
 RL Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK007731; BAB25219.1;  
 DR EMBL: BC003445; AA03445.1;  
 DR MGI: 1916430; 181003BD15Rik.  
 DR InterPro: IPR000520; Exonuclease.  
 DR Pfam: PF00929; Exonuclease; 1.  
 DR SMART: SM00479; EXOIII; 1.  
 SQ SEQUENCE 205 AA; 23613 MW; D3FFP9C7014B730E CRC64;  
 Query Match 64.3%; Score 18; DB 11; Length 205;  
 Best Local Similarity 40.0%; Pred. No. 43;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Oy 3 KKKXXXXXXLD 12  
 Db 157 KKAASHRALD 166  
 RESULT 14  
 Q8X4Z7 PRELIMINARY; PRT; 212 AA.  
 AC Q8X4Z7;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DE Putative tet operon regulator.  
 DE YCDC OR Z1512 OR ECS1259.  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=83334;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-0157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE-21074935; PubMed-11206551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grobleck E.J., Davis N.W., Lam A., Dimalanta E.T., Potamousis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.,  
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
 RL Nature 409:529-533(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-0157:H7 / RIMD 0509952;  
 RX MEDLINE-21156231; PubMed-11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
 RA Kubara S., Shiba T., Hattori M., Shinozawa H.,  
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli  
 O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RL DNA Res. 8:11-22(2001).  
 DR EMBL: AE005301; AAG555629.1; -

DR EMBL: AP002554; BAB34682.1; -  
 DR InterPro: IPR001647; HTH\_Tetr.  
 DR Pfam: PF00440; tetr. 1.  
 DR PRINTS: PR00455; HHTTTR.  
 KW Complete proteome.  
 SQ SEQUENCE 212 AA; 23703 MW; 5DMA7B85C960FDEC CRC64;  
 Query Match 64.3%; Score 18; DB 16; Length 212;  
 Best Local Similarity 40.0%; Pred. No. 45;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Oy 3 KKKXXXXXXLD 12  
 Db 20 KKAILLSALD 29  
 RESULT 15  
 Q96609 PRELIMINARY; PRT; 215 AA.  
 AC Q96609;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE Surface antigen atell.  
 OS Entamoeba histolytica.  
 OC Eukaryota; Entamoebidae; Entamoeba.  
 OX NCBI\_TaxID=5759;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-HM1-IMSS;  
 RX MEDLINE-98084495; PubMed-9423879;  
 RA Mai Z., Samuelson J.,  
 RT "A new gene family (arief) encodes asparagine-rich Entamoeba  
 RT histolytica antigens, which resemble the anebic vaccine candidate  
 RT serine-rich E. histolytica protein.";  
 RL Infect. Immun. 66:353-355(1998).  
 DR EMBL: AF093575; AAC72364.1; -  
 SQ SEQUENCE 215 AA; 23641 MW; 00B59B87816A5ECD CRC64;  
 Query Match 64.3%; Score 18; DB 5; Length 215;  
 Best Local Similarity 40.0%; Pred. No. 45;  
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Oy 3 KKKXXXXXXLD 12  
 Db 50 KKSSNSELD 59

Search completed: June 5, 2003, 15:55:45  
 Job time : 34.9091 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:47:43 ; Search time 40.7273 Seconds  
(without alignments)  
39.261 Million cell updates/sec

Title: US-09-150-947F-18  
Perfect score: 17  
Sequence: 1 XXXXXXXXXXXXD 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /SID2/gcgdata/geneseq/emb1/AA1980.DAT:\*  
2: /SID2/gcgdata/geneseq/emb1/AA1981.DAT:\*  
3: /SID2/gcgdata/geneseq/emb1/AA1982.DAT:\*  
4: /SID2/gcgdata/geneseq/emb1/AA1983.DAT:\*  
5: /SID2/gcgdata/geneseq/emb1/AA1984.DAT:\*  
6: /SID2/gcgdata/geneseq/emb1/AA1985.DAT:\*  
7: /SID2/gcgdata/geneseq/emb1/AA1986.DAT:\*  
8: /SID2/gcgdata/geneseq/emb1/AA1987.DAT:\*  
9: /SID2/gcgdata/geneseq/emb1/AA1988.DAT:\*  
10: /SID2/gcgdata/geneseq/emb1/AA1989.DAT:\*  
11: /SID2/gcgdata/geneseq/emb1/AA1990.DAT:\*  
12: /SID2/gcgdata/geneseq/emb1/AA1991.DAT:\*  
13: /SID2/gcgdata/geneseq/emb1/AA1992.DAT:\*  
14: /SID2/gcgdata/geneseq/emb1/AA1993.DAT:\*  
15: /SID2/gcgdata/geneseq/emb1/AA1994.DAT:\*  
16: /SID2/gcgdata/geneseq/emb1/AA1995.DAT:\*  
17: /SID2/gcgdata/geneseq/emb1/AA1996.DAT:\*  
18: /SID2/gcgdata/geneseq/emb1/AA1997.DAT:\*  
19: /SID2/gcgdata/geneseq/emb1/AA1998.DAT:\*  
20: /SID2/gcgdata/geneseq/emb1/AA1999.DAT:\*  
21: /SID2/gcgdata/geneseq/emb1/AA2000.DAT:\*  
22: /SID2/gcgdata/geneseq/emb1/AA2001.DAT:\*  
23: /SID2/gcgdata/geneseq/emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6	35.3	2	AA10185	Sequence of synth
2	6	35.3	2	AA37936	Trypsin modulating
3	6	35.3	3	AA40819	Sequence of iragme
4	6	35.3	3	AA04607	Antiviral agent.
5	6	35.3	3	AA00718	Core repeat of cel
6	6	35.3	3	AA25615	Harvey Ras inhibit
7	6	35.3	3	AA36707	Adhesion formation
8	6	35.3	3	AA30728	Igf-mast cell reac
9	6	35.3	3	AA30753	Igf-mast cell reac
10	6	35.3	3	AA53144	RGD peptide deriva

11	6	35.3	3	AA44666
12	6	35.3	3	AA48960
13	6	35.3	3	AA82967
14	6	35.3	3	AA70472
15	6	35.3	3	AA64556
16	6	35.3	3	AA64722
17	6	35.3	3	AA11094
18	6	35.3	3	AA99823
19	6	35.3	3	AA99825
20	6	35.3	3	AA99827
21	6	35.3	3	AA39130
22	6	35.3	3	AA31143
23	6	35.3	3	AA25187
24	6	35.3	3	AA25193
25	6	35.3	3	AA25173
26	6	35.3	3	AA64738
27	6	35.3	3	AA71120
28	6	35.3	3	AA48589
29	6	35.3	3	AA56245
30	6	35.3	3	AA56225
31	6	35.3	3	AA56186
32	6	35.3	3	AA56197
33	6	35.3	3	AA50604
34	6	35.3	3	AA50605
35	6	35.3	3	AA43493
36	6	35.3	3	AA43494
37	6	35.3	3	AA39889
38	6	35.3	3	AA30593
39	6	35.3	3	AA32860
40	6	35.3	3	AA32860
41	6	35.3	3	AA32872
42	6	35.3	3	AA13741
43	6	35.3	3	AA88037
44	6	35.3	3	AA37920
45	6	35.3	3	AA37922

# ALIGNMENTS

RESULT 1	AA10185	AA10185 standard; Protein; 2 AA.
ID	AA10185	
AC	AA10185	
XX		
DT	16-AUG-2002 (updated)	
DT	14-AUG-1992 (first entry)	
DE	Sequence of synthetic polypeptide for the prodn. of aspartame.	
DE	Aspartyl-phenylalanine methyl ester; artificial sweetener;	
KW	aspartame.	
XX		
OS	Synthetic.	
XX		
PN	EP36258-A.	
XX		
PD	23-SEP-1981.	
XX		
PF	02-MAR-1981; 81EP-0300857.	
XX		
PR	14-MAR-1980; 80US-0130462.	
XX		
PA	(CETU-) CETUS CORP.	
XX		
PI	Rose JE, White TJ, Bahl CP;	
XX		
DR	WPI; 1981-72133D/40.	
XX		
DR	N-PSDB; AA10059.	
XX		
PT	Aspartame prodn. from synthetic polypeptide - produced by cloned microorganism	

Platelet aggregati  
NL4-3 truncated GA  
Non-RGD, non-YISGR  
Cancer metastasis  
RF-1 peptide 1 fro  
HPR3 peptide deriv  
Platelet-targetin  
Active domain from  
Mutant chemokine a  
Human IL-16 5' end  
Platelet-targeting  
LDV-peptide capabl  
IDA-peptide capabl  
RGD-peptide capabl  
Angiotensin II pep  
Peptide AII(1-3) u  
Integrin receptor  
Anti-Inflammatory  
Anti-Inflammatory  
Anti-Inflammatory  
Anti-Inflammatory  
Resin bound cyclic  
Resin bound cyclic  
Linker for dual av  
Asparaginyl endope  
Amino acid sequenc  
Fibronectin protei  
Fibronectin protei  
RGD peptide. Unid  
Peptide used in th  
Trypsin modulating  
Trypsin modulating

XX Disclosure; page 5; 18pp; English.

XX The inventors claim a method for the prodn. of aspartame by first  
CC synthesizing double-stranded DNA in which a coding strand has  
CC alternating codons for Asp and the Phe in sufficient number to  
CC produce a polypeptide which is stable in predetermined host  
CC microorganisms. The DNA strand is then inserted into a cloning  
CC vehicle so that the resulting chimera directs the synthesis of the  
CC Asp-Phe protein. This protein is then cleaved with CNBr or trypsin  
CC to release the peptide which, after benzylation, can be fragmented  
CC by digestion with chymotrypsin.  
CC (Updated on 16-AUG-2002 to add missing OS field.)

XX  
SQ Sequence 2 AA;

Query Match 35.3%; Score 6; DB 2; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.4e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12  
1 D 1

DB 1 D 1

RESULT 2  
AAB37936  
ID AAB37936 standard; peptide; 2 AA.  
XX  
AC AAB37936;  
XX  
DT 27-FEB-2001 (first entry)  
XX  
DE Trypsin modulating oostatic factor (TMOF) peptide SEQ ID 41.  
XX  
XX  
XX Trypsin modulating oostatic factor; TMOF; pesticide; pest control;  
KW digestive enzyme synthesis inhibition; mosquito larvae; coleopteran;  
KW lepidopteran; dipteran; blood-sucking insect; Diptera; Nematocera;  
KW Colicidae; Culicinae; Corethrinae; Ceratopogonidae; Simuliidae.  
OS Synthetic.  
XX  
XX WO200063233-A2.  
PN  
XX  
PD 26-OCT-2000.  
XX  
PF 18-APR-2000; 2000WO-US10235.  
XX  
PR 21-APR-1999; 99US-0295996.  
XX  
XX (UFL ) UNIV FLORIDA RES FOUND INC.  
PA (UYNC-) UNIV NORTH CAROLINA STATE.  
XX  
XX Borovsky D, Linderman RJ;  
XX  
XX WPI: 2000-687156/67.  
DR  
XX  
XX Novel trypsin modulating oostatic factor compound peptides, useful for  
PT pest control including pests of agricultural crops -  
XX  
XX  
PS Claim 3; Page 33; 50pp; English.  
XX  
XX This invention relates to a trypsin modulating oostatic factor (TMOF)  
CC compound. The compound specifically does not include the TMOF analogue  
CC peptides represented by AAB37942 - AAB37949. Sequences AAB37900 -  
CC AAB37941 are TMOF analogues which may be included in the composition. The  
CC TMOF compound has pesticidal activity and is an inhibitor of pest  
CC digestive enzyme synthesis. The TMOF compound is useful for controlling  
CC pests such as mosquito larvae, coleopterans, lepidopterans, dipterans or  
CC blood-sucking insects of order Diptera, suborder Nematocera, family  
CC Colicidae or subfamily Culicinae, Corethrinae, Ceratopogonidae and  
CC Simuliidae.

SQ Sequence 2 AA;

Query Match 35.3%; Score 6; DB 21; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.4e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12  
1 D 2

DB 2 D 2

RESULT 3  
AAP40819  
ID AAP40819 standard; peptide; 3 AA.  
XX  
AC AAP40819;  
XX  
DT 03-AUG-1992 (first entry)  
XX  
DE Sequence of fragment D, corresp. to residues 25-27 of human  
DE pancreatic growth hormone releasing factor (somatostatin) (hpgRF).  
XX  
XX Hormone; dwarfism; therapy; retarded growth;  
KW anabolic protein deficiency; growth promoter; lactation.  
XX  
XX Homo sapiens.  
XX  
XX  
XX Key location/Qualifiers  
FH Modified-site 1  
FT /Label= Boc-D(OBzl)  
FT /note= "Boc-tertbutyloxycarbonyl (carbamate);  
FT OBzl=ester benzyllique"  
FT Modified-site 3  
FT /Label= M-NH-NH2  
XX  
XX  
XX AU8424774-A.  
PN  
XX  
XX 30-AUG-1984.  
PD  
XX  
XX 20-FEB-1984; 84AU-0400343.  
PF  
XX  
XX 29-NOV-1983; 83FR-0019058.  
PR  
XX 21-FEB-1983; 83FR-0002781.  
XX  
XX (SNFI ) SANOFI SA.  
XX  
XX Diaz J, Demarne H, Roncucci R, Schmelck PH;  
XX  
XX WPI: 1984-256760/42.  
DR  
XX  
XX Synthesis of hpgRF in liquid phase reactions - with use of new  
PT peptide fragments  
XX  
XX  
XX Claim 4; Page 67; 76pp; French.  
XX  
XX The inventors claim hpgRF fragments used for the synthesis of hpgRF.  
CC Using the method somatostatin can be obt'd. on a large scale with  
CC good yield and good purity. The prod. is used in man for the  
CC treatment of dwarfism and retarded growth and for anabolic protein  
CC deficiencies. In animals it is useful for promoting wt. growth for  
CC increased prodn. lactation etc. This index is based on EP-122818,  
CC which is the equivalent of AU8424774.  
XX  
XX  
SQ Sequence 3 AA;

Query Match 35.3%; Score 6; DB 5; Length 3;  
Best Local Similarity 100.0%; Pred. No. 7.4e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12  
1 D 1

DB 1 D 1

RESULT 4  
 AAR04607 standard; peptide; 3 AA.  
 ID AAR04607 standard; peptide; 3 AA.  
 AC AAR04607;  
 DT 05-SEP-1990 (first entry)  
 DE Antiviral agent.  
 DE Antiviral agent.  
 KM Antiviral; M2; poliovirus; polio; hepatitis.  
 OS Synthetic.  
 PN JP02078631-A.  
 PD 19-MAR-1990.  
 PF 14-SEP-1988; 88JP-0228843.  
 PR 14-SEP-1988; 88JP-0228843.  
 PS (NHA) NIPPON MINING KK.  
 DR WPI; 1990-129060/17.  
 PT Antiviral agent contg. tripeptide (unit) -  
 PT of basic aminoacid, then alanine, glycine or sarcosine, and  
 PT acide aminoacid, effective against virus with protein-terminated DNA  
 or RNA.  
 PS Disclosure; 4pp; Japanese.  
 CC Peptide is effective against inhibiting propagation of DNA or RNA  
 CC bonded, protein containing viruses eg. Poliovirus, Hepatitis virus.  
 SQ Sequence 3 AA;  
 Query Match 35.3%; Score 6; DB 11; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 12 D 12  
 DB 3 D 3

RESULT 5  
 AAR00718 standard; peptide; 3 AA.  
 ID AAR00718 standard; peptide; 3 AA.  
 AC AAR00718;  
 DT 29-MAY-1990 (first entry)  
 DE Core repeat of cell-adhesive protein.  
 DE Core repeat of cell-adhesive protein.  
 KM Cell adhesion; anti-metastatic agent; immunomodulation; core repeat.  
 OS  
 PN EP347931-A.  
 PD 27-DEC-1989.  
 PF 23-JUN-1989; 89EP-0111468.  
 PR 24-JUN-1988; 88JP-0156133.  
 PS (AZUMU) AZUMA I.  
 DR Saki I, Nishi N, Azuma I, Tokura S;  
 WPI; 1990-001305/01.

Polypeptide with repeated sequences of cell adhesion protein used  
 as anti-metastatic agent for cancer and agonist or antagonist of  
 cell-adhesion proteins  
 Claim 2; page 14; 16pp; English.  
 Peptide core is repeated 2-20 times to form a cell-adhesive protein of  
 mol. wt. 1,500-5,000. The protein is an (ant)agonist of cell-adhesive  
 proteins such as fibronectin. It has high antimetastatic activity  
 against cancer and can be used in immunomodulation, wound healing,  
 platelet aggregation inhibition and alleviation of neuro-disorders.  
 See also AAR00722.  
 Sequence 3 AA;  
 Query Match 35.3%; Score 6; DB 11; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 12 D 12  
 DB 3 D 3

RESULT 6  
 AAR25615 standard; peptide; 3 AA.  
 ID AAR25615 standard; peptide; 3 AA.  
 AC AAR25615;  
 DT 20-JAN-1993 (first entry)  
 DE Harvey Ras inhibitor - Ras 30-32.  
 KM GTPase; activating protein; GAP; oncogenesis; cancer; truncated;  
 KM colorectal; exocrine pancreatic; myeloid leukemia; adenyl-acylase.  
 OS Synthetic.  
 PN EP496162-A.  
 PD 29-JUL-1992.  
 PF 16-DEC-1991; 91EP-0311658.  
 PR 24-DEC-1990; 90US-0632891.  
 PS (MERI) MERCK & CO INC.  
 DR Freidinger RM, Garsky VM, Gibbs JB, Schaber MD, Sigal IS;  
 WPI; 1992-251427/31.  
 Synthetic peptide inhibitors of Ras-GAP interaction - as  
 antitumour agents against e.g. colorectal carcinoma, exocrine  
 pancreatic carcinoma and myeloid leukaemia  
 Claim 3; Page 7; 31pp; English.  
 The peptide corresponds to a C- and N-terminal truncated version of  
 the prod. of the Harvey ras oncogene from residues 17-44, i.e. from  
 residues 30-32. The peptide is capable of binding to GTPase  
 activating protein (GAP) thus inhibiting Ras activation by GAP and  
 inhibiting Ras-dependent cell transformation. Derivs. of the Ras  
 peptide are also provided, comprising C-terminal truncations, N-  
 terminal truncations, or modifications of specific amino acids, e.g.  
 the most potent Ras inhibitor is Ras 17-32. Oncogenic forms of Ras  
 occur in various human cancers, e.g. colorectal carcinoma, exocrine  
 pancreatic carcinoma and myeloid leukaemias. Ras complexed with GTP  
 can bind GAP and this is thought to be the important step in the  
 transformation activity of Ras. The peptide is also capable of  
 inhibiting GAP as well as the interaction of Ras with yeast

CC adenylyl acylase. See also AAR25602-29.  
XX  
SQ Sequence 3 AA;

Query Match  
Best Local Similarity 35.3%; Score 6; DB 13; Length 3;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
DB 1 D 1

RESULT 7  
AAR36707  
ID AAR36707 standard; peptide; 3 AA.

AC AAR36707;

DT 26-MAY-1993 (first entry)

DE Adhesion formation prevention RGD-contg. peptide.

KW Tissue repair; peritoneum; surgery; post-surgically; inhibition;  
platelet aggregation; cardiovascular; orthopedic; thoracic;  
ophthalmic; CNS; use.

OS Synthetic.

PN MO9308818-A.

PD 13-MAY-1993.

PF 06-NOV-1992; 92WO-US09494.

PR 07-NOV-1991; 91US-0789231.

PA (UTSC-) UNIV SOUTHERN CALIFORNIA.

PI Dizerega GS, Rodgers KE;

DR WPI, 1993-167381/20.

PT Prevention of adhesion formation, partic. post-surgically - comprises  
administering a RGD-contg. peptide for a time sufficient to permit  
tissue repair

PS Example: Page 18; 22pp; English.

CC The sequence is that of an RGD-contg. peptide which is used in a  
method for prevention of adhesion formation for a time sufficient  
to permit tissue repair. The method is used for minimizing or  
preventing adhesion formation, partic. in the peritoneum following  
surgery, but also for e.g. cardiovascular, orthopedic, thoracic,  
ophthalmic, CNS and other uses. In addn., the peptide inhibits  
platelet aggregation and does not induce inflammation or trauma  
at the site of administration.

CC  
CC  
SQ Sequence 3 AA;

Query Match  
Best Local Similarity 35.3%; Score 6; DB 14; Length 3;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
DB 3 D 3

RESULT 8  
AAR30728  
ID AAR30728 standard; peptide; 3 AA.  
XX

AC AAR30728;  
XX  
XX 20-MAY-1993 (first entry)  
DT  
XX

DE IGE-mast cell reaction preventing peptide intermediate.

KW Basophil; protective group; fragment condensing.

XX  
XX Synthetic.

OS  
FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-t-butoxycarbonyl-beta-benzyl-L-aspartyl"

PN JP04360898-A.

PD 14-DEC-1992.

PF 04-JUN-1991; 91JP-0159394.

PR 04-JUN-1991; 91JP-0159394.

PA (TAMA) TANABE SEIYAKU CO.

DR WPI; 1993-032710/04.

DT L-Aspartyl-L-seryl-L-aspartyl-L-prolyl-L-arginine prepn.

PT Involves fragment-condensing tripeptide with amino-protected  
di-peptide to form penta-peptide

PS Claim 7; Page 2; 8pp; Japanese.

CC The prepn. of DSPR is claimed and involves e.g. fragment-condensing  
C-protected tripeptide with N-protected dipeptide to form pentapeptide  
and then removing the protective gps.  
CC N-t-butoxycarbonyl-beta-benzyl-L-aspartyl-L-prolyl-L-arginine is a  
new cpd."

CC  
SQ Sequence 3 AA;

Query Match  
Best Local Similarity 35.3%; Score 6; DB 14; Length 3;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
DB 1 D 1

RESULT 9  
AAR30753  
ID AAR30753 standard; peptide; 3 AA.

AC AAR30753;

DT 26-MAY-1993 (first entry)

DE IGE-mast cell reaction inhibitory peptide prepn. peptide.

XX Physiologically active; basophil; tripeptide.

XX  
XX Synthetic.

OS Key Location/Qualifiers

FH Modified-site 1 /note= "N-t-butoxycarbonyl-beta-benzyl-L-Asp"

FT Modified-site 5 /note= "N(G)-nitro-L-Arg benzyl ester"

PN JP04360899-A.

PD 14-DEC-1992.



AC AAR48960;  
 XX  
 DT 12-SEP-1994 (first entry)  
 XX  
 DE NL4-3 truncated GAG C-terminal peptide.  
 XX  
 KM HIV-1; HXB2: antisense; sequence inversion; antisense virus; infection;  
 KM naturally occurring virus; NOV; translation; replication; infectivity;  
 KM hepatitis B; HIV-2; SIV; flip-over PCR.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9403596-A.  
 XX  
 PD 17-FEB-1994.  
 XX  
 PF 30-JUL-1993; 93WO-US071179.  
 XX  
 PR 30-JUL-1992; 92US-0921104.  
 XX  
 PA (UYHA-) UNIV HAMAM.  
 XX  
 PI Hu W, Wang J;  
 XX  
 DR WPI: 1994-065685/08.  
 DR N-PSDB: AAO57688.  
 XX  
 DT New antisense viruses and anti-sense-ribozyme viruses - used for  
 PT treating or preventing viral infections, partic. HIV-1, HIV-2 or  
 PT SIV infection  
 XX  
 PS Disclosure: Page 108; 167pp; English.  
 XX  
 CC This sequence is encoded by a PCR fragment of NL4-3 and represents the  
 CC C-terminal peptide fragment of the truncated GAG protein. The DNA  
 CC encoding this fragment was ligated into ClaI/SalI digested px and the  
 CC corresponding plasmid was used to produce the antisense virus of the  
 CC invention. Antisense or truncated RNAs expressed by these viruses  
 CC bind to the mRNAs expressed by the naturally occurring viruses (NOVs)  
 CC and prevent the NOV from being translated into proteins, thereby  
 CC preventing the NOV from replicating. The antisense viruses maintain  
 CC the infectivity of the NOV, allowing antisense RNAs to reach the  
 CC mRNAs of the natural viruses. Antisense viruses such as these may be  
 CC used for treating or preventing a viral infection, particularly HIV-1,  
 CC HIV-2 or SIV infection or hepatitis B infection.  
 CC  
 SQ Sequence 3 AA;  
 Query Match 35.3%; Score 6; DB 15; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 12 D 12  
 DB 2 D 2  
 RESULT 13  
 AAR82907  
 ID AAR82907 standard; peptide; 3 AA.  
 XX  
 AC AAR82907;  
 XX  
 DT 20-DEC-1995 (first entry)  
 XX  
 DE Non-RGD, non-YISGR cancer metastasis inhibitory peptide #1.  
 XX  
 KM Cancer metastasis; adhesive peptide; core sequence; dextran; cancer;  
 KM water soluble polysaccharide; metastasis; wound; immunogenicity.  
 OS Synthetic.  
 XX  
 PN JP07089999-A.

XX  
 PD 04-APR-1995.  
 XX  
 PF 17-SEP-1993; 93JP-0254779.  
 XX  
 PR 17-SEP-1993; 93JP-0254779.  
 XX  
 PA (JAPG ) NIPPON ZEON KK.  
 XX  
 DR WPI: 1995-167254/22.  
 XX  
 PT Cancer metastasis inhibitive peptide derivs. - useful for inhibition  
 PT of cancer metastasis, healing of wounds and regulation of  
 PT immunogenicity.  
 XX  
 PS Disclosure: Page 3; 6pp; Japanese.  
 XX  
 CC The peptides AAR70472-90 and AAR82907-24 are peptide derivatives which  
 CC inhibit cancer metastasis. They are composed of an adhesive peptide  
 CC with a core sequence selected from: RGD (AAR70472-85), YIGSR  
 CC (AAR70486-90) or other sequence (AAR82907-24), linked to a water soluble  
 CC polysaccharide, preferably a water soluble dextran, at the C-terminus.  
 CC The peptides are useful in inhibiting cancer metastasis, healing wounds  
 CC and the regulation of immunogenicity.  
 CC  
 SQ Sequence 3 AA;  
 Query Match 35.3%; Score 6; DB 16; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 12 D 12  
 DB 2 D 2  
 RESULT 14  
 AAR70472  
 ID AAR70472 standard; peptide; 3 AA.  
 XX  
 AC AAR70472;  
 XX  
 DT 20-DEC-1995 (first entry)  
 XX  
 DE Cancer metastasis inhibitory peptide core RGD sequence.  
 XX  
 KM Cancer metastasis; adhesive peptide; core sequence; dextran; cancer;  
 KM water soluble polysaccharide; metastasis; wound; immunogenicity.  
 OS Synthetic.  
 XX  
 PN JP07089999-A.  
 XX  
 PD 04-APR-1995.  
 XX  
 PF 17-SEP-1993; 93JP-0254779.  
 XX  
 PR 17-SEP-1993; 93JP-0254779.  
 XX  
 PA (JAPG ) NIPPON ZEON KK.  
 XX  
 DR WPI: 1995-167254/22.  
 XX  
 PT Cancer metastasis inhibitive peptide derivs. - useful for inhibition  
 PT of cancer metastasis, healing of wounds and regulation of  
 PT immunogenicity.  
 XX  
 PS Disclosure: Page 2; 6pp; Japanese.  
 XX  
 CC The peptides AAR70472-90 and AAR82907-24 are peptide derivatives which  
 CC inhibit cancer metastasis. They are composed of an adhesive peptide  
 CC with a core sequence selected from: RGD (AAR70472-85), YIGSR  
 CC (AAR70486-90) or other sequence (AAR82907-24), linked to a water soluble



CC polysaccharide, preferably a water soluble dextran, at the C-terminus.  
 CC The peptides are useful in inhibiting cancer metastasis, healing wounds  
 CC and the regulation of immunogenicity.

XX Sequence 3 AA:

Query Match 35.3%; Score 6; DB 16; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12

Db 3 D 3

RESULT 15

ID AAR64556 standard; Peptide; 3 AA.

AC AAR64556;

DT 01-SEP-1995 (first entry)

DE RF-1 peptide 1 from respiratory syncytial virus.

KW antiviral activity; DP-178; DP-107; diagnostic; HIV-1LAI;  
 human immunodeficiency virus; transmembrane protein; gp41;

KW alpha helix; leucine zipper; DP-185; respiratory syncytial virus;  
 RSV.

XX Synthetic.

OS Key Location/Qualifiers

FT Modified-site 1 /note- "optionally has an amino, acetyl,  
 9-fluorenylmethoxy-carbonyl, hydrophobic or  
 macromolecular carrier gp. attached"

FT Modified-site 3 /note- "optionally has a carboxyl, amide, hydrophobic  
 or macromolecular carrier gp. attached"

FT Modified-site 3 /note- "optionally has a carboxyl, amide, hydrophobic  
 or macromolecular carrier gp. attached"

XX WO9428920-A.

XX 22-DEC-1994.

PF 07-JUN-1994; 94NO-US05739.

PR 07-JUN-1993; 93US-0073028.

XX (UYDU-) UNIV DUKE.

PI Barney SO, Bolognesi DP, Lambert DM, Matthews TJ;  
 Petteway SR, Wild CT;

DR MPI; 1995-036105/05.

PT Computer search generated synthetic peptides - are inhibitors of  
 HIV transmission

XX Claim 14; Page 137; 182pp; English.

CC AAR64556-589 are peptide derivatives of a 37 mer RF-1 peptide derived  
 CC from respiratory syncytial virus (RSV) (AAR64590) which have been  
 CC truncated at the carboxy terminus. The peptides are DP-178 like  
 CC peptides. DP-178 corresponds to amino acids 638 to 673 of the HIV-1  
 CC isolate LAI transmembrane protein gp41. It forms a putative alpha  
 CC helix at the C-terminal end of the gp41 ectodomain, and complexes  
 CC with DP-107 (corresponds to amino acids 558-595) which contains a  
 CC leucine zipper motif. The peptides complex via non-covalent  
 CC protein-protein interactions. The peptide derivatives were  
 CC identified by a computer assisted peptide sequence search. The  
 CC antiviral activity of this peptide is not stated in the  
 CC specification.

XX Sequence 3 AA:

Query Match 35.3%; Score 6; DB 16; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12

Db 3 D 3

Search completed: June 5, 2003, 16:00:25  
 Job time : 40.7273 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw.model

Run on: June 5, 2003, 15:23:33 ; Search time 10.5455 Seconds  
(without alignments)  
33,481 Million cell updates/sec

Title: US-09-150-947F-18  
Perfect score: 17  
Sequence: 1 xxxxxxxxxxxx 12

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

## Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/laa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/laa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/laa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/laa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/laa/PCTus.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/laa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	35.3	2	1	US-07-820-154A-15
2	6	35.3	2	1	US-07-820-154A-23
3	6	35.3	2	1	US-07-820-154A-33
4	6	35.3	2	1	US-07-820-154A-39
5	6	35.3	2	1	US-08-358-160-171
6	6	35.3	2	1	US-08-153-799-20
7	6	35.3	2	2	US-08-097-554A-15
8	6	35.3	2	2	US-08-097-554A-23
9	6	35.3	2	2	US-08-097-554A-33
10	6	35.3	2	2	US-08-097-554A-39
11	6	35.3	2	2	US-08-539-432-15
12	6	35.3	2	2	US-08-539-432-16
13	6	35.3	2	2	US-08-539-432-17
14	6	35.3	2	2	US-09-327-424-2
15	6	35.3	2	3	US-08-480-640A-15
16	6	35.3	2	3	US-08-480-640A-23
17	6	35.3	2	3	US-08-480-640A-33
18	6	35.3	2	3	US-08-480-640A-39
19	6	35.3	2	3	US-08-295-802-15
20	6	35.3	2	3	US-08-295-802-23
21	6	35.3	2	3	US-08-295-802-33
22	6	35.3	2	3	US-08-295-802-39
23	6	35.3	2	4	US-09-326-335-2
24	6	35.3	2	4	US-08-488-237A-15
25	6	35.3	2	4	US-08-488-237A-23
26	6	35.3	2	4	US-08-488-237A-33
27	6	35.3	2	4	US-08-488-237A-39

28	6	35.3	2	4	US-08-375-992A-15	Sequence 15, App1
29	6	35.3	2	4	US-08-375-992A-23	Sequence 23, App1
30	6	35.3	2	4	US-08-375-992A-33	Sequence 33, App1
31	6	35.3	2	4	US-08-375-992A-39	Sequence 39, App1
32	6	35.3	2	4	US-09-295-996B-23	Sequence 23, App1
33	6	35.3	2	4	US-09-295-996B-35	Sequence 35, App1
34	6	35.3	2	4	US-09-295-996B-58	Sequence 58, App1
35	6	35.3	2	5	PCT-US93-00324-15	Sequence 15, App1
36	6	35.3	2	5	PCT-US93-00324-23	Sequence 23, App1
37	6	35.3	2	5	PCT-US93-00324-33	Sequence 33, App1
38	6	35.3	2	5	PCT-US93-00324-39	Sequence 39, App1
39	6	35.3	3	1	US-07-748-943-2	Sequence 2, App1
40	6	35.3	3	1	US-07-748-943-1	Sequence 1, App1
41	6	35.3	3	1	US-08-169-524-6	Sequence 6, App1
42	6	35.3	3	1	US-08-081-539-117	Sequence 117, App
43	6	35.3	3	1	US-08-251-027-1	Sequence 1, App1
44	6	35.3	3	1	US-08-251-027-2	Sequence 2, App1
45	6	35.3	3	1	US-08-251-027-3	Sequence 3, App1

## ALIGNMENTS

RESULT 1  
US-07-820-154A-15  
Sequence 15, Application US/07820154A  
Patent No. 5382425  
GENERAL INFORMATION:  
APPLICANT: Cochran Ph.D., Mark D  
APPLICANT: Junker M.S., David E  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/820,154A  
FILING DATE: 19920113  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)977-9550  
TELEFAX: (212)664-0525  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-820-154A-15  
Query Match 35.3%, Score 6, DB 1, Length 2;  
Best Local Similarity 100.0%, Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 12 D 12  
Db 1 D 1  
RESULT 2  
US-07-820-154A-23

Sequence 23, Application US/07820154A  
Patent No. 5382425  
GENERAL INFORMATION:  
APPLICANT: Cochran Ph.D., Mark D  
APPLICANT: Junker M.S., David E  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/820,154A  
FILING DATE: 19920113  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)977-9550  
TELEFAX: (212)664-0525  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-820-154A-23

Query Match 35.3%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12  
|  
Db 1 D 1

RESULT 3  
US-07-820-154A-33  
Sequence 33, Application US/07820154A  
Patent No. 5382425  
GENERAL INFORMATION:  
APPLICANT: Cochran Ph.D., Mark D  
APPLICANT: Junker M.S., David E  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/820,154A  
FILING DATE: 19920113  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)977-9550  
TELEFAX: (212)664-0525  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-820-154A-33

Query Match 35.3%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12  
|  
Db 1 D 1

RESULT 4  
US-07-820-154A-39  
Sequence 39, Application US/07820154A  
Patent No. 5382425  
GENERAL INFORMATION:  
APPLICANT: Cochran Ph.D., Mark D  
APPLICANT: Junker M.S., David E  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/820,154A  
FILING DATE: 19920113  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)977-9550  
TELEFAX: (212)664-0525  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-820-154A-39

Query Match 35.3%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12  
|  
Db 1 D 1

RESULT 5  
US-08-358-160-171  
Sequence 171, Application US/08358160  
Patent No. 5663143

GENERAL INFORMATION:  
APPLICANT: Let, Arthur C.  
APPLICANT: LADNER, Robert C.  
APPLICANT: GUTERMAN, Sonia K.  
APPLICANT: ROBERTS, Bruce L.  
APPLICANT: MARKLAND, William  
APPLICANT: KENT, Rachel B.  
TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ  
TITLE OF INVENTION: DOMAINS THAT INHIBIT HUMAN NEUTROPHIL ELASTASE  
NUMBER OF SEQUENCES: 234  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W. Suite 300  
CITY: Washington  
STATE: District of Columbia  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/358,160  
FILING DATE: 16-DEC-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/133,031  
FILING DATE: 13-OCT-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/009,319  
FILING DATE: 26-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/664,989  
FILING DATE: 01-MAR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/487,063  
FILING DATE: 02-MAR-1990  
APPLICATION NUMBER: US 07/240,160  
FILING DATE: 02-SEP-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Cooper, Iver P.  
REGISTRATION NUMBER: 28,005  
REFERENCE/DOCKET NUMBER: LEY-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 171:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-358-160-171

Query Match 35.3%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12  
DB 1 D 1

RESULT 6  
US-08-153-799-20  
Sequence 20, Application US/08153799  
Patent No. 5766883  
GENERAL INFORMATION:  
APPLICANT: Ballance, David J

APPLICANT: Goodey, Andrew R  
TITLE OF INVENTION: Polypeptides  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: R Hain Swope, BOC Health Care Inc  
STREET: 100 Mountain Avenue  
CITY: Murray Hill  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07974  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/153,799  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/847975  
FILING DATE: 06-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 8909916.2  
FILING DATE: 29-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB90/00650  
FILING DATE: 26-APR-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/775952  
FILING DATE: 29-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Swope, R Hain  
REGISTRATION NUMBER: 24864  
REFERENCE/DOCKET NUMBER: 92H832  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 665 2400  
TELEFAX: (908) 771 6159  
TELEX: 219484  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-153-799-20

Query Match 35.3%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12  
DB 1 D 1

RESULT 7  
US-08-097-554A-15  
Sequence 15, Application US/08097554A  
Patent No. 5869312  
GENERAL INFORMATION:  
APPLICANT: Cochran Ph.D., Mark D  
APPLICANT: Junker M.S., David E  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 112  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,554A  
FILING DATE: July 22, 1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)977-9550  
TELEFAX: (212)664-0525  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-097-554A-15

Query Match 35.3%; Score 6; DB 2; Length 2;  
Best Local Similarity 100.0%; Pred.No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
1 D 1

RESULT 8  
US-08-097-554A-23  
Sequence 23, Application US/08097554A  
Patent No. 5869312  
GENERAL INFORMATION:  
APPLICANT: Cochran Ph.D., Mark D  
APPLICANT: Junker M.S., David E  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 112  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,554A  
FILING DATE: July 22, 1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)977-9550  
TELEFAX: (212)664-0525  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-097-554A-23

Query Match 35.3%; Score 6; DB 2; Length 2;  
Best Local Similarity 100.0%; Pred.No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
1 D 1

RESULT 9  
US-08-097-554A-33  
Sequence 33, Application US/08097554A  
Patent No. 5869312  
GENERAL INFORMATION:  
APPLICANT: Cochran Ph.D., Mark D  
APPLICANT: Junker M.S., David E  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 112  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,554A  
FILING DATE: July 22, 1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)977-9550  
TELEFAX: (212)664-0525  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-097-554A-33

Query Match 35.3%; Score 6; DB 2; Length 2;  
Best Local Similarity 100.0%; Pred.No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
1 D 1

RESULT 10  
US-08-097-554A-39  
Sequence 39, Application US/08097554A  
Patent No. 5869312  
GENERAL INFORMATION:  
APPLICANT: Cochran Ph.D., Mark D  
APPLICANT: Junker M.S., David E  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 112  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,554A  
FILING DATE: JULY 22, 1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)977-9550  
TELEFAX: (212)664-0525  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-097-534A-39

Query Match 35.3%; Score 6; DB 2; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
DB 1 D 1

RESULT 11  
US-08-539-432-15  
Sequence 15, Application US/08539432  
Patent No. 5872210  
GENERAL INFORMATION:  
APPLICANT: MEDABALIMI, JOHN L.  
TITLE OF INVENTION: TRANSFRAME INHIBITORY  
TITLE OF INVENTION: ELEMENT OF VIRAL  
TITLE OF INVENTION: PROTEASE  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10054  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/539,432  
FILING DATE: 05-OCT-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4194  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 751-6849  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2  
TYPE: Amino Acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Linear  
US-08-539-432-15

Query Match 35.3%; Score 6; DB 2; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
DB 1 D 1

RESULT 12  
US-08-539-432-16  
Sequence 16, Application US/08539432  
Patent No. 5872210  
GENERAL INFORMATION:  
APPLICANT: MEDABALIMI, JOHN L.  
TITLE OF INVENTION: TRANSFRAME INHIBITORY  
TITLE OF INVENTION: ELEMENT OF VIRAL  
TITLE OF INVENTION: PROTEASE  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10054  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/539,432  
FILING DATE: 05-OCT-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4194  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 751-6849  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2  
TYPE: Amino Acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Linear  
US-08-539-432-16

Query Match 35.3%; Score 6; DB 2; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
DB 1 D 1

RESULT 13  
US-08-539-432-17  
Sequence 17, Application US/08539432  
Patent No. 5872210  
GENERAL INFORMATION:  
APPLICANT: MEDABALIMI, JOHN L.  
TITLE OF INVENTION: TRANSFRAME INHIBITORY  
TITLE OF INVENTION: ELEMENT OF VIRAL  
TITLE OF INVENTION: PROTEASE  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10054

COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/539,432  
FILING DATE: 05-OCT-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4194  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2  
TYPE: Amino Acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Linear  
US-08-539-432-17

Query Match 35.3%; Score 6; DB 2; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
|  
DB 1 D 1

RESULT 14  
US-09-327-424-2  
; Sequence 2, Application US/09327424  
; Patent No. 6027903  
; GENERAL INFORMATION:  
; APPLICANT: Eyre, David R  
; TITLE OF INVENTION: KIT FOR DETECTING ANALYTE INDICATIVE OF TYPE I COLLAGEN  
; TITLE OF INVENTION: RESORPTION IN VIVO (as amended)  
; FILE REFERENCE: WROS-1-14019  
; CURRENT APPLICATION NUMBER: US/09/327,424  
; CURRENT FILING DATE: 1999-06-07  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Human  
; FEATURE:  
; OTHER INFORMATION: N-telopeptide sequence of type II collagen  
US-09-327-424-2

Query Match 35.3%; Score 6; DB 3; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
|  
DB 2 D 2

RESULT 15  
US-08-480-640A-15  
; Sequence 15, Application US/08480640A  
; Patent No. 6033904  
; GENERAL INFORMATION:  
; APPLICANT: Cochran, Mark D.  
; APPLICANT: Junker, David E.  
; TITLE OF INVENTION: Recombinant Swinepox Virus  
; NUMBER OF SEQUENCES: 225  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: John P. White  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,640A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
REGISTRATION NUMBER: 28,678  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-480-640A-15

Query Match 35.3%; Score 6; DB 3; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
|  
DB 1 D 1

Search completed: June 5, 2003, 15:32:44  
Job time : 10.5455 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: June 5, 2003, 15:31:01; Search time 14.5455 Seconds

(without alignments)  
85.173 Million cell updates/secTitle: US-09-150-947f-18  
Perfect score: 17  
Sequence: 1 xxxxxxxxxxxx 12Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database: Published Applications\_AA.\*

- 1: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubppa/PT1\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubppa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubppa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubppa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubppa/PT05\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubppa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubppa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubppa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubppa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6	35.3	2	9	US-10-187-339-6
2	6	35.3	2	9	US-10-206-699-296
3	6	35.3	2	10	US-09-982-172-8
4	6	35.3	2	10	US-09-982-172-105
5	6	35.3	3	9	US-09-900-936-11
6	6	35.3	3	9	US-09-840-277-31
7	6	35.3	3	9	US-09-747-408-21
8	6	35.3	3	9	US-10-014-162-37
9	6	35.3	3	9	US-10-039-876A-3
10	6	35.3	3	9	US-10-039-876A-4
11	6	35.3	3	9	US-10-015-879-58
12	6	35.3	3	9	US-10-226-956-316
13	6	35.3	3	9	US-10-226-956-317
14	6	35.3	3	9	US-10-226-956-318
15	6	35.3	3	9	US-10-226-956-319
16	6	35.3	3	9	US-10-128-198-1
17	6	35.3	3	9	US-10-237-850-81
18	6	35.3	3	9	US-10-032-818-6
19	6	35.3	3	9	US-10-281-652-11

20	6	35.3	3	9	US-10-283-838-13	Sequence 13, Appl
21	6	35.3	3	9	US-10-163-587A-19	Sequence 19, Appl
22	6	35.3	3	9	US-10-141-531-15	Sequence 15, Appl
23	6	35.3	3	9	US-10-141-531-16	Sequence 16, Appl
24	6	35.3	3	9	US-10-141-531-17	Sequence 17, Appl
25	6	35.3	3	9	US-10-141-531-18	Sequence 18, Appl
26	6	35.3	3	9	US-10-141-531-19	Sequence 19, Appl
27	6	35.3	3	9	US-10-141-531-20	Sequence 20, Appl
28	6	35.3	3	10	US-09-096-749A-113	Sequence 113, App
29	6	35.3	3	10	US-09-866-824A-9	Sequence 9, Appl
30	6	35.3	3	10	US-09-792-2008-19	Sequence 19, Appl
31	6	35.3	3	10	US-09-771-192-11	Sequence 11, Appl
32	6	35.3	3	10	US-09-853-918-1	Sequence 1, Appl
33	6	35.3	3	10	US-09-853-918-6	Sequence 6, Appl
34	6	35.3	3	10	US-09-882-667-14	Sequence 14, Appl
35	6	35.3	3	10	US-09-867-847-30	Sequence 30, Appl
36	6	35.3	3	10	US-09-822-261-7	Sequence 7, Appl
37	6	35.3	3	10	US-09-312-762A-12	Sequence 12, Appl
38	6	35.3	3	10	US-09-982-172-184	Sequence 184, App
39	6	35.3	3	10	US-09-837-697A-11	Sequence 11, Appl
40	6	35.3	3	12	US-10-003-674A-6	Sequence 6, Appl
41	6	35.3	4	1	US-08-464-363-35	Sequence 35, Appl
42	6	35.3	4	1	US-08-610-220A-2	Sequence 2, Appl
43	6	35.3	4	1	US-08-610-220A-7	Sequence 7, Appl
44	6	35.3	4	1	US-08-610-220A-8	Sequence 8, Appl
45	6	35.3	4	8	US-08-765-244-16	Sequence 16, Appl

## ALIGNMENTS

RESULT 1  
US-10-187-339-6  
Sequence 6, Application US/10187339  
Publication No. US20030084478A1  
GENERAL INFORMATION:  
APPLICANT: Ream, Walt et al.,  
TITLE OF INVENTION: Plants Having Enhanced Gail Resistance and  
FILE REFERENCE: 53629  
CURRENT APPLICATION NUMBER: US/10/187,339  
CURRENT FILING DATE: 2002-06-28  
PRIOR APPLICATION NUMBER: US/09/434,837  
PRIOR FILING DATE: 1999-11-04  
PRIOR APPLICATION NUMBER: 60/107,185  
PRIOR FILING DATE: 1998-11-05  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 6  
LENGTH: 2  
TYPE: PRT  
ORGANISM: Agrobacterium tumefaciens  
US-10-187-339-6

Query Match 35.3%; Score 6; DB 9; Length 2;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0;  
Gaps 0;

QY 12 D 12  
Db 2 D 2

RESULT 2  
US-10-206-699-296  
Sequence 296, Application US/10206699  
Publication No. US20030100510A1  
GENERAL INFORMATION:  
APPLICANT: Sundaramoorthy, M.  
TITLE OF INVENTION: Crystallized structure of Type IV Collagen NCI Domain Hexamer  
FILE REFERENCE: MHB 01-1017  
CURRENT APPLICATION NUMBER: US/10/206,699

```
;; CURRENT FILING DATE: 2002-07-26
;; PRIOR APPLICATION NUMBER: US 60/308,523
;; PRIOR FILING DATE: 2001-07-27
;; PRIOR APPLICATION NUMBER: US 60/351,289
;; PRIOR FILING DATE: 2001-10-29
;; PRIOR APPLICATION NUMBER: US 60/366,854
;; PRIOR FILING DATE: 2002-03-22
;; PRIOR APPLICATION NUMBER: US 60/385,362
;; PRIOR FILING DATE: 2002-06-03
;; NUMBER OF SEQ ID NOS: 307
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 296
;; LENGTH: 2
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-206-699-296

Query Match
Best Local Similarity 35.3%; Score 6; DB 9; Length 2;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12
DB 2 D 2

RESULT 3
US-09-982-172-8
;; Sequence 8, Application US/09982172
;; Patent No. US20020137119A1
;; GENERAL INFORMATION:
;; APPLICANT: Emil Israel Katz
;; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
;; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERAL
;; FILE REFERENCE: 01/22283
;; CURRENT APPLICATION NUMBER: US/09/982,172
;; CURRENT FILING DATE: 2001-10-19
;; NUMBER OF SEQ ID NOS: 253
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 8
;; LENGTH: 2
;; TYPE: PRT
;; ORGANISM: Artificial sequence
;; FEATURE:
;; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-8

Query Match
Best Local Similarity 35.3%; Score 6; DB 10; Length 2;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12
DB 1 D 1

RESULT 4
US-09-982-172-105
;; Sequence 105, Application US/09982172
;; Patent No. US20020137119A1
;; GENERAL INFORMATION:
;; APPLICANT: Emil Israel Katz
;; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
;; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERAL
;; FILE REFERENCE: 01/22283
;; CURRENT APPLICATION NUMBER: US/09/982,172
;; CURRENT FILING DATE: 2001-10-19
;; NUMBER OF SEQ ID NOS: 253
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 105
;; LENGTH: 2
```

```
;; TYPE: PRT
;; ORGANISM: Artificial sequence
;; FEATURE:
;; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-105

Query Match
Best Local Similarity 35.3%; Score 6; DB 10; Length 2;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12
DB 1 D 1

RESULT 5
US-09-900-936-11
;; Sequence 11, Application US/09900936
;; Patent No. US20020165141A1
;; GENERAL INFORMATION:
;; APPLICANT: Rodgers, Kathleen
;; APPLICANT: dizerega, Gere
;; TITLE OF INVENTION: Methods for Promoting Dendritic Cell Proliferation
;; TITLE OF INVENTION: or Differentiation
;; FILE REFERENCE: 00-506-A
;; CURRENT APPLICATION NUMBER: US/09/900,936
;; CURRENT FILING DATE: 2001-07-09
;; NUMBER OF SEQ ID NOS: 50
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 11
;; LENGTH: 3
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: AII (1-3)
US-09-900-936-11

Query Match
Best Local Similarity 35.3%; Score 6; DB 9; Length 3;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12
DB 1 D 1

RESULT 6
US-09-840-277-31
;; Sequence 31, Application US/09840277
;; Patent No. US20020168363A1
;; GENERAL INFORMATION:
;; APPLICANT: FEIGE, ULRICH
;; APPLICANT: KOHNO, TADAHITO
;; APPLICANT: LACEY, DAVID LEE
;; APPLICANT: BOONE, THOMAS CHARLES
;; TITLE OF INVENTION: INTEGRIN/ADHESION ANTAGONISTS
;; FILE REFERENCE: A-688A
;; CURRENT APPLICATION NUMBER: US/09/840,277
;; CURRENT FILING DATE: 2001-08-14
;; PRIOR APPLICATION NUMBER: 60/198,919
;; PRIOR FILING DATE: 2000-04-21
;; PRIOR APPLICATION NUMBER: 60/201,394
;; PRIOR FILING DATE: 2000-05-03
;; NUMBER OF SEQ ID NOS: 135
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 31
;; LENGTH: 3
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Integrin antagonist peptide
US-09-840-277-31
```

Query Match 35.3%; Score 6; DB 9; Length 3;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12  
DB 3 D 3

RESULT 7  
US-09-747-408-21  
; Sequence 21, Application US/09747408  
; Publication No. US20030003141A1  
; GENERAL INFORMATION:  
; APPLICANT: Green, Allan M.  
; APPLICANT: Gervais, Francine  
; TITLE OF INVENTION: Compounds And Methods For Modulating  
; FILE REFERENCE: NBI-088  
; CURRENT APPLICATION NUMBER: US/09/747,408  
; CURRENT FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/171,877  
; PRIOR FILING DATE: 1999-12-23  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 21  
; LENGTH: 3  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-747-408-21

Query Match 35.3%; Score 6; DB 9; Length 3;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12  
DB 1 D 1

RESULT 8  
US-10-014-162-37  
; Sequence 37, Application US/10014162  
; Publication No. US20030032096A1  
; GENERAL INFORMATION:  
; APPLICANT: Usdin, Ted B.  
; APPLICANT: Hoare, Samuel R.J.  
; TITLE OF INVENTION: PARATHYROID HORMONE RECEPTOR LIGANDS  
; FILE REFERENCE: NIH175.001C1  
; CURRENT APPLICATION NUMBER: US/10/014,162  
; CURRENT FILING DATE: 2001-12-11  
; PRIOR APPLICATION NUMBER: PCT/US00/1677  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 60/139335  
; PRIOR FILING DATE: 1999-06-15  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 37  
; LENGTH: 3  
; TYPE: PRT  
; ORGANISM: Bos taurus  
US-10-014-162-37

Query Match 35.3%; Score 6; DB 9; Length 3;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12  
DB 1 D 1

RESULT 9

US-10-039-876A-3  
; Sequence 3, Application US/10039876A  
; Publication No. US20030032792A1  
; GENERAL INFORMATION:  
; APPLICANT: Conklin, Darrell C.  
; APPLICANT: Blumberg, Hal  
; TITLE OF INVENTION: A HUMAN 2-19 PROTEIN HOMOLOGUE, 2219A  
; FILE REFERENCE: 97-63C1  
; CURRENT APPLICATION NUMBER: US/10/039,876A  
; CURRENT FILING DATE: 2001-10-26  
; PRIOR APPLICATION NUMBER: US 60/061,712  
; PRIOR FILING DATE: 1997-10-06  
; PRIOR APPLICATION NUMBER: US 09/167,513  
; PRIOR FILING DATE: 1998-10-06  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 3  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Motif 1, corresponding to residues 127 to 129 of  
; OTHER INFORMATION: SEQ ID NO:2  
US-10-039-876A-3

Query Match 35.3%; Score 6; DB 9; Length 3;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12  
DB 2 D 2

RESULT 10  
US-10-039-876A-4  
; Sequence 4, Application US/10039876A  
; Publication No. US20030032792A1  
; GENERAL INFORMATION:  
; APPLICANT: Conklin, Darrell C.  
; APPLICANT: Blumberg, Hal  
; TITLE OF INVENTION: A HUMAN 2-19 PROTEIN HOMOLOGUE, 2219A  
; FILE REFERENCE: 97-63C1  
; CURRENT APPLICATION NUMBER: US/10/039,876A  
; CURRENT FILING DATE: 2001-10-26  
; PRIOR APPLICATION NUMBER: US 60/061,712  
; PRIOR FILING DATE: 1997-10-06  
; PRIOR APPLICATION NUMBER: US 09/167,513  
; PRIOR FILING DATE: 1998-10-06  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 3  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Motif 2, corresponding to residues 156 to 158 of  
; OTHER INFORMATION: SEQ ID NO:2  
US-10-039-876A-4

Query Match 35.3%; Score 6; DB 9; Length 3;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12  
DB 2 D 2

RESULT 11  
US-10-015-979-58  
; Sequence 58, Application US/10015979  
; Publication No. US20030036050A1

```
; GENERAL INFORMATION:
; APPLICANT: Quax, Wilhelmus J.
; APPLICANT: Verhaert, Raymond M.D.
; APPLICANT: Beekwilder, Martinus J.
; APPLICANT: Aehle, Wolfgang
; TITLE OF INVENTION: Enzyme Selection
; FILE REFERENCE: 2183-520705
; CURRENT APPLICATION NUMBER: US/10/015,979
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: PCT/NL00/00399
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/138,443
; PRIOR FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: CB4 domain of mutant 13
; NAME/KEY: SITE
; LOCATION: (1)..(3)
; OTHER INFORMATION:
US-10-015-979-58
```

```
Query Match          35.3%; Score 6; DB 9; Length 3;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      12 D 12
DB      3 D 3
```

```
RESULT 12
US-10-226-956-316
; Sequence 316, Application US/10226956
; Publication No. US20030060399A1
; GENERAL INFORMATION:
; APPLICANT: Brophy, Colleen
; APPLICANT: Komalavilas, Padmini
; APPLICANT: Panitch, Alyssa
; APPLICANT: Joshi, Lokesh
; APPLICANT: Seal, Brandon L.
; TITLE OF INVENTION: REAGENTS AND METHODS FOR SMOOTH MUSCLE THERAPIES
; FILE REFERENCE: ASU-1061-US
; CURRENT APPLICATION NUMBER: US/10/226,956
; CURRENT FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: 60/314,535
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 320
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 316
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-10-226-956-316
```

```
Query Match          35.3%; Score 6; DB 9; Length 3;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      12 D 12
DB      1 D 1
```

```
RESULT 13
US-10-226-956-317
```

```
; Sequence 317, Application US/10226956
; Publication No. US20030060399A1
; GENERAL INFORMATION:
; APPLICANT: Brophy, Colleen
; APPLICANT: Komalavilas, Padmini
; APPLICANT: Panitch, Alyssa
; APPLICANT: Joshi, Lokesh
; APPLICANT: Seal, Brandon L.
; TITLE OF INVENTION: REAGENTS AND METHODS FOR SMOOTH MUSCLE THERAPIES
; FILE REFERENCE: ASU-1061-US
; CURRENT APPLICATION NUMBER: US/10/226,956
; CURRENT FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: 60/314,535
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 320
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 317
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-10-226-956-317
```

```
Query Match          35.3%; Score 6; DB 9; Length 3;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      12 D 12
DB      1 D 1
```

```
RESULT 14
US-10-226-956-318
; Sequence 318, Application US/10226956
; Publication No. US20030060399A1
; GENERAL INFORMATION:
; APPLICANT: Brophy, Colleen
; APPLICANT: Komalavilas, Padmini
; APPLICANT: Panitch, Alyssa
; APPLICANT: Joshi, Lokesh
; APPLICANT: Seal, Brandon L.
; TITLE OF INVENTION: REAGENTS AND METHODS FOR SMOOTH MUSCLE THERAPIES
; FILE REFERENCE: ASU-1061-US
; CURRENT APPLICATION NUMBER: US/10/226,956
; CURRENT FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: 60/314,535
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 320
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 318
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-10-226-956-318
```

```
Query Match          35.3%; Score 6; DB 9; Length 3;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      12 D 12
DB      1 D 1
```

```
RESULT 15
US-10-226-956-319
; Sequence 319, Application US/10226956
; Publication No. US20030060399A1
; GENERAL INFORMATION:
```

; APPLICANT: Brophy, Colleen  
 ; APPLICANT: Komalavilas, Padmini  
 ; APPLICANT: Panitch, Alyssa  
 ; APPLICANT: Joshi, Lokesh  
 ; APPLICANT: Seal, Brandon L.  
 ; TITLE OF INVENTION: REAGENTS AND METHODS FOR SMOOTH MUSCLE THERAPIES  
 ; FILE REFERENCE: ASU-1061-US  
 ; CURRENT APPLICATION NUMBER: US/10/226,956  
 ; CURRENT FILING DATE: 2002-08-23  
 ; PRIOR APPLICATION NUMBER: 60/314,535  
 ; PRIOR FILING DATE: 2001-08-23  
 ; NUMBER OF SEQ ID NOS: 320  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 319  
 ; LENGTH: 3  
 ; TYPE: PRT  
 ; ORGANISM: Artificial sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Synthetic peptide  
 US-10-226-956-319

Query Match 35.3%; Score 6; DB 9; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 12 D 12  
 Db 1 D 1

Search completed: June 5, 2003, 15:49:05  
 Job time : 14.5455 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:24:01 ; Search time 11.8182 seconds  
(without alignments)  
97.613 Million cell updates/sec

Title: US-09-150-947f-18  
Perfect score: 17  
Sequence: 1 XXXXXXXXXXXXD 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: PIR1:  
2: PIR2:  
3: PIR3:  
4: PIR4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	35.3	4	2	140697
2	6	35.3	4	2	141890
3	6	35.3	4	2	D41654
4	6	35.3	4	2	A32480
5	6	35.3	4	2	IS1049
6	6	35.3	4	2	PT0271
7	6	35.3	4	2	B53284
8	6	35.3	4	2	PT0696
9	6	35.3	4	2	PT0711
10	6	35.3	4	2	A26209
11	6	35.3	5	2	A32516
12	6	35.3	5	2	C23751
13	6	35.3	5	2	A26830
14	6	35.3	5	2	A32014
15	6	35.3	5	2	B31836
16	6	35.3	5	2	B60274
17	6	35.3	5	2	D60274
18	6	35.3	5	2	S70615
19	6	35.3	5	2	PO0689
20	6	35.3	5	2	PT0267
21	6	35.3	5	2	PT0281
22	6	35.3	5	2	PT0308
23	6	35.3	5	2	PT0596
24	6	35.3	5	2	PT0513
25	6	35.3	5	2	PT0600
26	6	35.3	5	2	PT0729
27	6	35.3	5	2	PT0624
28	6	35.3	5	2	PT0601
29	6	35.3	5	2	PT0672

30	6	35.3	5	2	PT0660	T-cell receptor be
31	6	35.3	5	2	PT0651	T-cell receptor be
32	6	35.3	5	2	PT0656	T-cell receptor be
33	6	35.3	5	2	PT0535	T-cell receptor be
34	6	35.3	5	2	PT0699	T-cell receptor be
35	6	35.3	5	2	PT0538	T-cell receptor be
36	6	35.3	5	2	PT0561	T-cell receptor be
37	6	35.3	5	2	PT0540	T-cell receptor be
38	6	35.3	5	2	PT0703	T-cell receptor be
39	6	35.3	5	2	PT0690	T-cell receptor be
40	6	35.3	5	2	PT0573	T-cell receptor be
41	6	35.3	5	2	PT0580	T-cell receptor be
42	6	35.3	5	2	PT0679	T-cell receptor be
43	6	35.3	5	2	S68326	blood cell protein
44	6	35.3	6	2	A61419	sarcosine dehydrog
45	6	35.3	6	2	B44510	hypothetical prote

## ALIGNMENTS

RESULT 1  
140697  
blotin A - Citrobacter freundii (fragment)  
C:Species: Citrobacter freundii  
C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 12-Aug-1996  
C:Accession: 140697  
R:Shiuan, D.; Campbell, A.  
Gene 67, 203-211, 1988  
A:Title: Transcriptional regulation and gene arrangement of Escherichia coli, Citrobacter  
A:Reference number: 140697; MUID:89006280; PMID:2971595  
A:Accession: 140697  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-4 <RES>  
A:Cross-references: GB:M21922; NID:9144434

Query Match  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
Db 3 D 3

RESULT 2  
141890  
protein D - Escherichia coli (fragment)  
C:Species: Escherichia coli  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 30-Sep-1993  
C:Accession: A41890  
R:Stettan, A.; Gebhardt, K.; Kristiansen, E.; Birkenland, N.K.; Lindqvist, B.H.  
J. Bacteriol. 174, 4094-4100, 1992  
A:Title: Escherichia coli K-12 and B contain functional bacteriophage P2 cgr genes.  
A:Reference number: A41890; MUID:92283767; PMID:1597424  
A:Accession: A41890  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-4 <SLF>  
A:Cross-references: GB:M81463

Query Match  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
Db 2 D 2

RESULT 3  
D41654

hypothetical protein (sodc 5' region) - Haemophilus parainfluenzae (fragment)  
 C:Species: Haemophilus parainfluenzae  
 C:Date: 12-Jun-1992 #sequence\_revision 12-Jun-1992 #text\_change 24-Feb-1995  
 C:Accession: D41654  
 R:Kroll, J.S.; Langford, P.R.; Loynds, B.M.  
 J. Bacteriol. 173, 7449-7457, 1991  
 A:Title: Copper-zinc superoxide dismutase of Haemophilus influenzae and Haemophilus para  
 A:Reference number: A41654; MUID:92041655; PMID:1938942  
 A:Accession: D41654  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-4 <KRO>

Query Match 35.3%; Score 6; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
 DB 1 D 1

## RESULT 4

A32480  
 A:Title: Achatin-I - giant African snail  
 N:Contains: Achatin-II  
 C:Species: Achatina fulica (giant African snail)  
 C:Date: 12-Oct-1989 #sequence\_revision 12-Oct-1989 #text\_change 17-Mar-1999  
 C:Accession: A32480  
 R:Kamatani, Y.; Minakata, H.; Kenny, P.T.M.; Iwashita, T.; Watanabe, K.; Funase, K.; Sur  
 Biochem. Biophys. Res. Commun. 160, 1015-1020, 1989  
 A:Title: Achatin-I, an endogenous neuroexcitatory tetrapeptide from Achatina fulica feru  
 A:Reference number: A32480; MUID:89273551; PMID:2597281  
 A:Accession: A32480  
 A:Molecule type: protein  
 A:Residues: 1-4 <KAM>  
 A:Note: stereochemistry of the active form confirmed by chemical synthesis  
 R:Shida, T.; In, Y.; Inoue, M.; Yasuda-Kamatani, Y.; Minakata, H.; Iwashita, T.; Nomoto  
 FEBS Lett. 307, 253-256, 1992  
 A:Title: Effect of the D-Phe(2) residue on molecular conformation of an endogenous neuro  
 (H-Gly-Phe-Ala-Asp-OH).  
 A:Reference number: A44691; MUID:92354723; PMID:1644179  
 A:Contents: annotation; X-ray crystallography, 0.85 angstroms  
 A:Note: Achatin-II has L-phenylalanine  
 C:Keywords: D-amino acid  
 F:2/Modified site: D-phenylalanine (Phe) #status experimental

Query Match 35.3%; Score 6; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
 DB 4 D 4

## RESULT 5

151049  
 A:Title: metallothionein-A - rainbow trout (fragment)  
 C:Species: Oncorhynchus mykiss (rainbow trout)  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
 C:Accession: I51049  
 R:Olsson, P.E.; Kling, P.; Erkel, L.J.; Kille, P.  
 Eur. J. Biochem. 230, 344-349, 1995  
 A:Title: Structural and functional analysis of the rainbow trout (Oncorhynchus mykiss) me  
 A:Reference number: I51049; MUID:95324545; PMID:7601121  
 A:Accession: I51049  
 A:Status: Preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-4 <OIS>  
 A:Cross-references: EMBL:X80181; NID:q1019799; PIDN:CA5646.1; PID:g4379328

Query Match 35.3%; Score 6; DB 2; Length 4;

Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
 DB 2 D 2

## RESULT 6

PT0271  
 A:Title: heavy chain CRD3 region (clone 3-103A) - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
 C:Accession: PT0271  
 R:Yamada, M.; Messerman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
 J. Exp. Med. 173, 395-407, 1991  
 A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity an  
 A:Reference number: PT0222; MUID:91108337; PMID:1899102  
 A:Accession: PT0271  
 A:Molecule type: DNA  
 A:Residues: 1-4 <YAM>  
 A:Experimental source: B lymphocyte  
 C:Keywords: heterotetramer; immunoglobulin

Query Match 35.3%; Score 6; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
 DB 1 D 1

## RESULT 7

B53284  
 A:Title: T-cell receptor beta 2 chain D region, Dbeta2 - rabbit  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 02-May-1994 #sequence\_revision 18-Nov-1994 #text\_change 05-Nov-1999  
 C:Accession: B53284  
 R:Harindranath, N.; Alexander, C.B.; Mage, R.G.  
 Mol. Immunol. 28, 881-888, 1991  
 A:Title: Evolutionarily conserved organization and sequences of germline diversity an  
 A:Reference number: A53284; MUID:91342695; PMID:1678859  
 A:Accession: B53284  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-4 <HAR>  
 A:Cross-references: GB:S60737; NID:g233916; PIDN:AA19518.1; PID:g233918  
 A:Note: sequence extracted from NCBI Backbone (NCBIN:60737, NCBI:60738)  
 C:Keywords: T-cell receptor

Query Match 35.3%; Score 6; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
 DB 1 D 1

## RESULT 8

PT0696  
 A:Title: T-cell receptor beta chain V-D-J region (100-2N) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
 C:Accession: PT0696; PT0612; PT0545; PT0692; PT0552; PT0696  
 R:Feeney, A.J.  
 J. Exp. Med. 174, 115-124, 1991  
 A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions  
 A:Reference number: PT0696; MUID:91277601; PMID:1711558  
 A:Accession: PT0696  
 A:Status: translation not shown  
 A:Molecule type: mRNA



A:Residues: 1-4 <FE3>  
 A:Experimental source: newborn thymus, strain BALB/c, 100-2N  
 A:Accession: PT0612  
 A:Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-4 <FE2>  
 A:Experimental source: newborn thymus, strain BALB/c, 111-1J  
 A:Accession: PT0545  
 A:Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-4 <FE3>  
 A:Experimental source: day 18 fetal thymus, strain BALB/c, 126-1AD  
 A:Accession: PT0692  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-4 <FE4>  
 A:Experimental source: day 18 fetal thymus, strain BALB/c, 140-1M  
 A:Accession: PT0552  
 A:Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-4 <FE5>  
 A:Experimental source: day 18 fetal thymus, strain BALB/c, 126-1CI  
 A:Accession: PT0696  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-4 <FE6>  
 A:Experimental source: newborn thymus, strain BALB/c, 135-1AA  
 C:Keywords: T-cell receptor

Query Match 35.3%; Score 6; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
 DB 4 D 4

## RESULT 9

PT0711  
 T-cell receptor beta chain V-D-J region (120-22) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
 C:Accession: PT0607; PT0674; PT0570; PT0711; PT0710  
 R:Feeney, A.J.  
 J. Exp. Med. 174, 115-124, 1991  
 A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
 A:Reference number: PT0509; MUID:91277601; PMID:1711558  
 A:Accession: PT0607  
 A:Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-4 <FE1>  
 A:Experimental source: newborn thymus, strain BALB/c, 120-2J  
 A:Accession: PT0674  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-4 <FE2>  
 A:Experimental source: day 18 fetal thymus, strain BALB/c, 140-1G  
 A:Accession: PT0678  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-4 <FE3>  
 A:Experimental source: day 18 fetal thymus, strain BALB/c, 154-1L  
 A:Accession: PT0570  
 A:Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-4 <FE4>  
 A:Experimental source: day 19 fetal thymus, strain BALB/c, 141-1I  
 A:Accession: PT0711  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-4 <FE5>  
 A:Experimental source: newborn thymus, strain BALB/c (clones 161-2AF and 161-2AF)

C:Keywords: T-cell receptor

Query Match 35.3%; Score 6; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
 DB 4 D 4

## RESULT 10

A26209  
 protein-glutamine gamma-glutamyltransferase (EC 2.3.2.13) - guinea pig (fragment)  
 C:Species: Cavia porcellus (guinea pig)  
 C:Date: 10-Sep-1987 #sequence\_revision 10-Sep-1987 #text\_change 03-Mar-1995  
 C:Accession: A26209  
 R:Connell, J.M.; Chung, S.I.; Whetzel, N.K.; Bradley, L.M.; Folk, J.E.  
 J. Biol. Chem. 246, 1093-1098, 1971  
 A:Title: Structural properties of guinea pig liver transglutaminase.  
 A:Reference number: A26209; MUID:71111415; PMID:5543674  
 A:Accession: A26209  
 A:Molecule type: protein  
 A:Residues: 1-4 <CON>  
 A:Experimental source: liver  
 C:Keywords: aminocyltransferase

Query Match 35.3%; Score 6; DB 2; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
 DB 3 D 3

## RESULT 11

A32516  
 cholecystokinin-5 - dog  
 N:Alternate names: CCK-5  
 C:Species: Canis lupus familiaris (dog)  
 C:Date: 18-Oct-1989 #sequence\_revision 18-Oct-1989 #text\_change 18-Aug-2000  
 C:Accession: A32516  
 R:Shively, J.; Reeve Jr., J.R.; Eysselein, V.E.; Ben-Avram, C.; Vigna, S.R.; Walsh, J.  
 Am. J. Physiol. 252, G272-G275, 1987  
 A:Title: CCK-5: sequence analysis of a small cholecystokinin from canine brain and in  
 A:Reference number: A32516; MUID:87133871; PMID:3826354  
 A:Accession: A32516  
 A:Molecule type: protein  
 A:Residues: 1-5 <SH1>  
 C:Comment: This peptide corresponds to the five carboxyl-terminal residues of cholecy  
 C:Superfamily: gastrin  
 C:Keywords: amidated carboxyl end; neuropeptide  
 F:5/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 35.3%; Score 6; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
 DB 4 D 4

## RESULT 12

C23751  
 spinal cord peptide SCP-6 - pig  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 28-Sep-1987 #sequence\_revision 28-Sep-1987 #text\_change 18-Aug-2000  
 C:Accession: C23751  
 R:Hsi, K.L.; Chen, R.L.; Chen, Z.G.; Zhang, H.L.; Lu, Y.A.; Guo, S.Y.; Wu, S.X.; Tsou  
 Arch. Biochem. Biophys. 240, 178-183, 1985  
 A:Reference number: A23751; MUID:85250425; PMID:4015098

A:Accession: C23751  
 A:Status: preliminary  
 A:Molecule type: Protein  
 A:Residues: 1-5 <HS1>  
 C:Superfamily: unassigned animal peptides

Query Match 35.3%; Score 6; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
 |  
 Db 1 D 1

RESULT 13  
 A26830  
 C:Species: Mus musculus (house mouse)  
 C:Date: 19-Nov-1988 #sequence\_revision 18-Aug-2000 #text\_change 18-Aug-2000  
 C:Accession: A26830  
 R:Reichelt, K.; Elgyo, K.; Edmonson, P.D.  
 Biochem. Biophys. Res. Commun. 146, 1493-1501, 1987  
 A:Title: Isolation and structure of an epidermal mitosis inhibiting pentapeptide.  
 A:Reference number: A26830; PMID:87298602; PMID:3619940  
 A:Accession: A26830  
 A:Molecule type: Protein  
 A:Residues: 1-5 <RE1>  
 C:Superfamily: unassigned animal peptides  
 C:Keywords: blocked amino end; pyroglutamic acid  
 F1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 35.3%; Score 6; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
 |  
 Db 3 D 3

RESULT 14  
 A32014  
 C:Species: Escherichia coli  
 C:Date: 22-Jun-1989 #sequence\_revision 22-Jun-1989 #text\_change 16-Feb-1997  
 C:Accession: A32014  
 R:Imamoto, S.; Yoshioke, Y.; Ohtsubo, E.  
 J. Bacteriol. 170, 2749-2757, 1988  
 A:Title: Identification and characterization of the products from the trax and trax gene  
 A:Reference number: A32014; PMID:88227859; PMID:2836369  
 A:Accession: A32014  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-5 <INA>  
 C:Genetics:  
 A:Genome: plasmid  
 C:Keywords: DNA binding

Query Match 35.3%; Score 6; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
 |  
 Db 3 D 3

RESULT 15  
 B31836  
 C:Species: Rickettsia rickettsii (fragment)  
 C:Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 28-May-1999

C:Accession: B31836  
 R:Anderson, B.E.; Baumstark, B.R.; Bellini, W.J.  
 J. Bacteriol. 170, 4493-4500, 1988  
 A:Title: Expression of the gene encoding the 17-kilodalton antigen from Rickettsia r  
 A:Reference number: A31885; PMID:89008059; PMID:3139629  
 A:Accession: B31836  
 A:Molecule type: DNA  
 A:Residues: 1-5 <AND>  
 A:Cross-references: GB:J03371; NID:9152455; PID:AA015030.1; PID:94262874

Query Match 35.3%; Score 6; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
 |  
 Db 2 D 2

Search completed: June 5, 2003, 15:34:01  
 Job time : 11.8182 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:34:07 ; Search time 7.81818 Seconds

(Without alignments)  
63.661 Million cell updates/sec

Title: US-09-150-947f-18

Perfect score: 17

Sequence: 1 XXXXXXXXXXXXD 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892.

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	6	35.3	3	1	LUXE_VIBFI
2	6	35.3	4	1	ACH1_ACHFU
3	6	35.3	4	1	OCPI_OCTMI
4	6	35.3	4	1	OCPI_OCTMI
5	6	35.3	5	1	BIOA_CITFR
6	6	35.3	5	1	RELI_LITRU
7	6	35.3	5	1	TRM3_ECOLI
8	6	35.3	5	1	UXA4_CHITR
9	6	35.3	6	1	ASP2_LACSN
10	6	35.3	6	1	TRP1_PSEPU
11	6	35.3	7	1	ALL4_CARMA
12	6	35.3	7	1	ALL7_CYPDO
13	6	35.3	7	1	CHOK_AICSP
14	6	35.3	7	1	FAR1_HELTI
15	6	35.3	7	1	FAR2_PROCT
16	6	35.3	7	1	FARB_CALVO
17	6	35.3	7	1	UF03_MOUSE
18	6	35.3	7	1	UF04_MOUSE
19	6	35.3	8	1	ACT1_THRAL
20	6	35.3	8	1	ACT1_CARMA
21	6	35.3	8	1	AKH_MELML
22	6	35.3	8	1	AL12_CARMA
23	6	35.3	8	1	AL18_CARMA
24	6	35.3	8	1	ALL4_CALVO
25	6	35.3	8	1	ALL5_CALVO
26	6	35.3	8	1	ALL5_CYPDO
27	6	35.3	8	1	ANG2_BOTJA
28	6	35.3	8	1	CCKN_MACEU
29	6	35.3	8	1	COXG_RAT
30	6	35.3	8	1	FAR1_PANRE
31	6	35.3	8	1	FAR3_HOMAM
32	6	35.3	8	1	FARB_CALVO
33	6	35.3	8	1	GLOR_HUMAN

34	6	35.3	8	1	LCK1_LEUMA	P21140 leucophaea
35	6	35.3	8	1	LCK2_LEUMA	P21141 leucophaea
36	6	35.3	8	1	LCK3_LEUMA	P21142 leucophaea
37	6	35.3	8	1	LCK4_LEUMA	P21143 leucophaea
38	6	35.3	8	1	LCK7_LEUMA	P19989 leucophaea
39	6	35.3	8	1	LCK8_LEUMA	P19990 leucophaea
40	6	35.3	8	1	LMT2_LOCTI	P22356 locusta mlg
41	6	35.3	8	1	ORMY_ORCLI	P82455 orconectes
42	6	35.3	8	1	PLP_BRANA	P81707 brassica na
43	6	35.3	8	1	UC26_MAIZE	P80632 zea mays tm
44	6	35.3	8	1	UPA1_HUMAN	P30087 homo sapien
45	6	35.3	9	1	COM_CONVE	P83047 conus ventr

## ALIGNMENTS

RESULT 1	LOXE_VIBFI	STANDARD;	PRT;	3 AA.
AC P24272;	LOXE_VIBFI	STANDARD;	PRT;	3 AA.
DT 01-MAR-1992 (Rel. 21, Created)				
DT 01-MAR-1992 (Rel. 21, Last sequence update)				
DT 01-OCT-1996 (Rel. 34, Last annotation update)				
DE Long-chain-fatty-acid--luciferin-component ligase (EC 6.2.1.19) (Acyl- protein synthetase) (Fragment).				
GN LOXE.				
OS Vibrio fischeri.				
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.				
OX NCBI_TaxID=568;				
RN [1]				
RP SEQUENCE FROM N.A.				
RX MEDLINE=91072226; Pubmed=2254256;				
RA Swartzman E., Kapoor S., Graham A.F., Meighen E.A.;				
RT "A new Vibrio fischeri lux gene precedes a bidirectional termination site for the lux operon."				
RL J. Bacteriol. 172:6797-6802(1990).				
CC -1- FUNCTION: ACYL-PROTEIN SYNTHETASE ACTIVATES TETRADECANOIC ACID. IT IS A COMPONENT OF THE FATTY ACID REDUCTASE COMPLEX RESPONSIBLE FOR CONVERTING TETRADECANOIC ACID TO THE ALDEHYDE WHICH SERVES AS SUBSTRATE IN THE LUCIFERASE-CATALYZED REACTION.				
CC -1- CATALYTIC ACTIVITY: ATP + an acid + protein -> AMP + diphosphate + an acyl-protein thioester.				
CC -1- PATHWAY: Bioluminescent fatty acid reduction system; second step.				
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC				
DR EMBL; M62812; ; NOT_ANNOTATED_CDS.				
FT LUMINESCENCE; Ligase.				
KW NON_TER				
SQ SEQUENCE 3 AA; 374 MW; 6AA3303000000000 CMC64;				
Query Match	35.38;	Score 6;	DB 1;	Length 3;
Best Local Similarity	100.00;	Pred. No. 0;		
Matches 1;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
OY 12 D 12				
DB 3 D 3				
RESULT 2				
ACH1_ACHFU	STANDARD;	PRT;	4 AA.	
ID ACH1_ACHFU				
AC P35904;				
DT 01-JUN-1994 (Rel. 29, Created)				
DT 01-JUN-1994 (Rel. 29, Last sequence update)				

DT 15-JUN-1998 (Rel. 36, last annotation update)  
 DE Achatin-I.  
 OS Achatina fulica (Giant African snail).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;  
 OC Achatinaceae; Achatinidae; Achatina.  
 OX NCBI\_TaxID=6530;  
 RN [1]  
 RP SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.  
 RC STRAIN-Perussac; TISSUE-Ganglion;  
 RX MEDLINE=89273551; PubMed=2597281;  
 RA Kametani Y., Minakata H., Kenny P.T.M., Iwashita T., Matanabe K.,  
 RA Funase K., Sun X.P., Yongsiri A., Kim K.H., Novales-Ll P.,  
 RA Novales E.T., Kanapi C.G., Takeuchi H., Nomoto K.,  
 RT "Achatin-I", an endogenous neuroexcitatory tetrapeptide from Achatina  
 RT fulica Ferussac containing a D-amino acid residue.";  
 RL Biochem. Biophys. Res. Commun. 160:1015-1020(1989).  
 RN [2]  
 RP CHARACTERIZATION.  
 RC STRAIN-Perussac; TISSUE-Heart atrium;  
 RX MEDLINE=91264856; PubMed=1675568;  
 RA Fujimoto K., Kubota I., Yasuda-Kamatani Y., Minakata H., Nomoto K.,  
 RA Yoshida M., Harada A., Muneoka Y., Kobayashi M.;  
 RT "Purification of achatin-I from the atria of the African giant snail,  
 RT Achatina fulica, and its possible function.";  
 RL Biochem. Biophys. Res. Commun. 177:847-853(1991).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY.  
 RX MEDLINE=93014529; PubMed=1399265;  
 RA Ishida T., In Y., Doi M., Inoue M., Yasuda-Kamatani Y., Minakata H.,  
 RA Iwashita T., Nomoto K.;  
 RT "Crystal structure and molecular conformation of achatin-I  
 RT (H-Gly-D-Phe-Ala-Asp-OH), an endogenous neuropeptide containing a  
 RT D-amino acid residue.";  
 RL Int. J. Pept. Protein Res. 39:258-264(1992).  
 CC -1- FUNCTION: NEUROEXCITATORY PEPTIDE; INCREASES THE IMPULSE FREQUENCY  
 CC AND PRODUCES A SPIKE BROADENING OF THE IDENTIFIED HEART EXCITATORY  
 CC NEURON (PON); ALSO ENHANCES THE AMPLITUDE AND FREQUENCY OF THE  
 CC HEART BEAT. HAS ALSO AN EFFECT ON SEVERAL OTHER MUSCLES.  
 DR PIR: A32480; A32480.  
 KW Hormone; D-amino acid.  
 FT MOD.RES 2  
 SQ SEQUENCE 4 AA; 408 MW; 6AADD9C810000000 CRC64;  
 QY  
 DB 12 D 12  
 4 D 4  
 Query Match 35.3%; Score 6; DB 1; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 RESULT 3  
 OCP1\_OCTMI STANDARD; PRT; 4 AA.  
 AC P58648:  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, last sequence update)  
 DT 15-JUN-2002 (Rel. 41, last annotation update)  
 DE Cardioactive peptides Ocp-1/Ocp-2.  
 OS Octopus minor (Octopus).  
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;  
 OC Incurtata; Octopodidae; Octopus.  
 OX NCBI\_TaxID=89766;  
 RN [1]  
 RP SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.  
 RC TISSUE-Brain;  
 RX MEDLINE=20336815; PubMed=10876044;  
 RA Iwakoshi E., Hisada M., Minakata H.;  
 RT "Cardioactive peptides isolated from the brain of a Japanese octopus,  
 RT Octopus minor.";  
 RL Peptides 21:623-630(2000).  
 CC -1- FUNCTION: Cardioactive; has both positive chronotropic and  
 CC inotropic effects on the heart. Ocp-4 is a 1000 time less  
 CC active than Ocp-3.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- PTM: Ocp-4 has D-Ser instead of L-Ser.  
 CC -1- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.  
 KW Hormone; D-amino acid.  
 FT MOD.RES 2  
 SQ SEQUENCE 4 AA; 463 MW; 6AB365B810000000 CRC64;  
 QY  
 DB 12 D 12  
 4 D 4  
 Query Match 35.3%; Score 6; DB 1; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC Inotropic effects on the heart. Ocp-2 is a 1000 time less  
 CC active than Ocp-1.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- PTM: Ocp-2 has L-Phe instead of D-Phe.  
 CC -1- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.  
 KW Hormone; D-amino acid.  
 FT MOD.RES 2  
 SQ SEQUENCE 4 AA; 394 MW; 6AA879C810000000 CRC64;  
 QY  
 DB 12 D 12  
 4 D 4  
 Query Match 35.3%; Score 6; DB 1; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 RESULT 4  
 OCP3\_OCTMI STANDARD; PRT; 4 AA.  
 AC P58649:  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, last sequence update)  
 DT 15-JUN-2002 (Rel. 41, last annotation update)  
 DE Cardioactive peptides Ocp-3/Ocp-4.  
 OS Octopus minor (Octopus).  
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;  
 OC Incurtata; Octopodidae; Octopus.  
 OX NCBI\_TaxID=89766;  
 RN [1]  
 RP SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.  
 RC TISSUE-Brain;  
 RX MEDLINE=20336815; PubMed=10876044;  
 RA Iwakoshi E., Hisada M., Minakata H.;  
 RT "Cardioactive peptides isolated from the brain of a Japanese octopus,  
 RT Octopus minor.";  
 RL Peptides 21:623-630(2000).  
 CC -1- FUNCTION: Cardioactive; has both positive chronotropic and  
 CC inotropic effects on the heart. Ocp-4 is a 1000 time less  
 CC active than Ocp-3.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- PTM: Ocp-4 has D-Ser instead of L-Ser.  
 CC -1- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.  
 KW Hormone; D-amino acid.  
 FT MOD.RES 2  
 SQ SEQUENCE 4 AA; 463 MW; 6AB365B810000000 CRC64;  
 QY  
 DB 12 D 12  
 4 D 4  
 Query Match 35.3%; Score 6; DB 1; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 RESULT 5  
 BIOA\_CITFR STANDARD; PRT; 5 AA.  
 AC P13071:  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, last sequence update)  
 DT 16-OCT-2001 (Rel. 40, last annotation update)  
 DE Adenosylmethionine-8-amino-7-oxononanoate aminotransferase  
 DE (EC 2.6.1.62) (7,8-diamino-pelargonic acid aminotransferase) (DAPA  
 DE aminotransferase) (Fragment).  
 GN BIOA.  
 OS Citrobacter freundii.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Citrobacter.  
 OX NCBI\_TaxID=546;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE:89006280; PubMed:2971595;  
 RA Shiuan D., Campbell A.;  
 RT "transcriptional regulation and gene arrangement of *Escherichia coli*,  
 RT *Citrobacter freundii* and *Salmonella typhimurium* biotin operons.";  
 RL Gene 67:203-211(1988).  
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 8-amino-7-oxononanoate -> S-adenosyl-4-methylthio-2-oxobutanolate + 7,8-diaminononanoate.  
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.  
 CC -1- PATHWAY: Biotin biosynthesis.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT AMINOTRANSFERASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M21922; NOT\_ANNOTATED\_CDS.  
 DR Interpro: IPR000934; AminoTrans\_3.  
 DR PROSITE: PS00600; AA\_TRANSFER\_CLASS\_3; PARTIAL.  
 KM Biotin biosynthesis; Transferase; Amino transferase;  
 KM Pyridoxal phosphate.  
 FT NON\_TER 5  
 SQ SEQUENCE 5 AA: 582 MW: 66AAB1B1A6F00000 CRC64;  
 QY Query Match 35.3%; Score 6; DB 1; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 12 D 12  
 Db 4 D 4  
 RESULT 6  
 RE1\_LITRU STANDARD; PRT; 5 AA.  
 AC P82070;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Rubellidin 1.1.  
 OS *Litoria rubella* (Desert tree frog).  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;  
 CC Pelodytidae; Litoria.  
 OX NCBI\_TaxID=104895;  
 RN [1]  
 RP SEQUENCE AND MASS SPECTROMETRY.  
 RC TISSUE: Skin secretion;  
 RA Steinbocker S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C., Tyler M.J., Wallace J.C.;  
 RT "The structure of new peptides from the Australian red tree frog '*Litoria rubella*', the skin peptide profile as a probe for the study of evolutionary trends of amphibians.";  
 RL Aust. J. Chem. 49:955-963(1996).  
 CC -1- FUNCTION: SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR ANTIBIOTIC ACTIVITY.  
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.  
 CC -1- MASS SPECTROMETRY: MS-598; METHOD-FAB.  
 KM Amphibian skin.  
 SQ SEQUENCE 5 AA: 598 MW: 6DD9C9CA82A00000 CRC64;  
 QY Query Match 35.3%; Score 6; DB 1; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12  
 Db 2 D 2  
 RESULT 7  
 TRM3\_ECOLI STANDARD; PRT; 5 AA.  
 AC P13973;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Trm protein (Fragment).  
 GN TRAM.  
 OS *Escherichia coli*.  
 OC Plasmid IncFII R100.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC *Escherichia*.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE:88227859; PubMed:2836369;  
 RA Inamoto S., Yoshioka Y., Ohtsubo E.;  
 RT "Identification and characterization of the products from the trmJ and trm genes of plasmid R100.";  
 RL J. Bacteriol. 170:2749-2757(1988).  
 CC -1- FUNCTION: TRANSFER GENE PROTEIN. IS INVOLVED IN THE CONJUGATION PROCESS OF BACTERIAL CELLS FOR THE EXCHANGE OF PLASMID DNA.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: TO TRAM PROTEIN OF OTHER PLASMIDS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M20941; NOT\_ANNOTATED\_CDS.  
 DR PIR: A32014; A32014.  
 KM Conjugation; Plasmid; DNA-binding.  
 FT NON\_TER 1  
 SQ SEQUENCE 5 AA: 634 MW: 6B1B1AA443500000 CRC64;  
 QY Query Match 35.3%; Score 6; DB 1; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 12 D 12  
 Db 3 D 3  
 RESULT 8  
 UX44\_CHLTR STANDARD; PRT; 5 AA.  
 AC P38005;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Unknown protein from 2D-page from elementary body (Fragment).  
 OS *Chlamydia trachomatis*.  
 CC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=813;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=L2/434/Bu;  
 RA Bini L., Santucci A., Magi B., Marzocchi B., Sanchez-Campillo M., Comanducci M., Christensen G., Birkelund S., Viretton E., Ratti G., Pallini V.;  
 RL Submitted (SEP-1994) to the 2D-GEL THE DETERMINED P1 OF THIS UNKNOWN  
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED P1 OF THIS UNKNOWN

CC PROTEIN IS: 4.5, ITS MW IS: 28 kDa.  
 DR Stena-2DPAGE; P38005; -.  
 FT NON\_TER 5  
 SQ SEQUENCE 5 AA; 474 MW; 75BAA865AA800000 CRC64;  
 Query Match  
 Best Local Similarity 35.3%; Score 6; DB 1; Length 5;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 12 D 12  
 DB 4 D 4  
 RESULT 9  
 ASP2\_LACSN STANDARD; PRT; 6 AA.  
 AC P82655;  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Acid shock protein 2 (Fragment).  
 OS Lactobacillus sanfranciscensis (Lactobacillus sanfrancisco).  
 CC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;  
 CC Lactobacillus.  
 CC NCBI\_Taxid=1625;  
 RN (1)  
 RP SEQUENCE.  
 RC STRAIN=CBI;  
 RA MEDLINE-21322712; PubMed-11429463;  
 RX De Angelis M., Bini L., Pallini V., Coccocelli P.S., Gobetti M.;  
 RT "The acid-stress response in Lactobacillus sanfranciscensis CBI.";  
 RL Microbiology 147:1863-1873(2001).  
 CC -I- INDUCTION: OVEREXPRESSED IN ACID ENVIRONMENTS.  
 FT NON\_TER 6  
 SQ SEQUENCE 6 AA; 778 MW; 6AA45B5B132A8000 CRC64;  
 Query Match  
 Best Local Similarity 35.3%; Score 6; DB 1; Length 6;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 12 D 12  
 DB 6 D 6  
 RESULT 10  
 TRPI\_PSEPU STANDARD; PRT; 6 AA.  
 AC P36414;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-FEB-1995 (Rel. 31, Last annotation update)  
 DE TRPA operon transcriptional activator (Fragment).  
 GN TRPI.  
 OS Pseudomonas putida.  
 CC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 CC Pseudomonas.  
 CC NCBI\_Taxid=303;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PRGI C15;  
 RX MEDLINE-89335826; PubMed-2503057;  
 RA Eberly L., Crawford I.P.;  
 RT "DNA sequence of the tryptophan synthase genes of Pseudomonas putida.";  
 RL Blochmle 71:521-531(1989).  
 CC -I- FUNCTION: ACTIVATES THE EXPRESSION OF THE TRPA GENES ENCODING THE TWO TRYPTOPHAN SYNTHASE SUBUNITS. IN THE ABSENCE OF THE INDUCER (INDOLEGLYCEROL PHOSPHATE), TRPI BINDS UPSTREAM OF THE TRAP OPERON, OVERLAPPING ITS OWN PROMOTER REGION.  
 CC -I- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL REGULATORS.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X13299; CAA31660.1; -.  
 DR InterPro: IPR000847; HTH\_LYSR.  
 DR PROSITE: PS00044; HTH\_LYSR\_FAMILY: PARTIAL.  
 KW Tryptophan biosynthesis; Transcription regulation; Activator;  
 FT DNA-binding.  
 FT NON\_TER 6  
 SQ SEQUENCE 6 AA; 683 MW; 77672AA1EDD6F000 CRC64;  
 Query Match  
 Best Local Similarity 35.3%; Score 6; DB 1; Length 6;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 12 D 12  
 DB 4 D 4  
 RESULT 11  
 ALI4\_CARMA STANDARD; PRT; 7 AA.  
 AC P81807;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Carcinus statin 4.  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;  
 CC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;  
 CC Brachyura; Eubrachyura; Portunioidea; Portunidae; Carcinus.  
 CC NCBI\_Taxid=6759;  
 RN (1)  
 RP SEQUENCE.  
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion.  
 RX MEDLINE-98121193; PubMed-9461285;  
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
 RA Thorpe A.;  
 RT Isolation and identification of multiple neuropeptides of the  
 RT allatostatin superfamily in the shore crab Carcinus maenas.";  
 RL Eur. J. Biochem 250:727-734(1997).  
 CC -I- FUNCTION: MAY ACT AS A NEUROTANSITTER OR NEUROMODULATOR.  
 CC -I- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Multigene family.  
 SQ SEQUENCE 7 AA; 783 MW; 672879CDB476AC0 CRC64;  
 Query Match  
 Best Local Similarity 35.3%; Score 6; DB 1; Length 7;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 12 D 12  
 DB 1 D 1  
 RESULT 12  
 ALI7\_CYPDPO STANDARD; PRT; 7 AA.  
 AC P82158;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Cydistatin 7.  
 OS Cydia pomonella (Coddling moth).  
 CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 CC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;

OC Dityrisia; Tortricidae; Tortricidae; Olethreutinae; Cydia.  
 OX NCBI\_TaxID=82600;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE-Larva;  
 RX MEDLINE=98054539; PubMed=9392829;  
 RA Dave H., Johnson A.H., Maestro J.-L., Scott A.G., Winstanley D.,  
 RT "Peptidic peptides of the allatostatin superfamily."  
 RL Peptides 18:1301-1309(1997).  
 CC -1- SIMILARITY: BELONGS TO THE ALLANOSTATIN FAMILY.  
 CC Neuropeptide; Amidation.  
 FT MOD\_RES 7  
 SQ SEQUENCE 7 AA; 873 MW; 672879CABB569350 CRC64;  
 QY Query Match 35.3%; Score 6; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 12 D 12  
 4 D 4

RESULT 13  
 CHOX\_ALCSP STANDARD; PRT; 7 AA.  
 ID ID  
 AC P16101;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 01-APR-1990 (Rel. 14, Last annotation update)  
 DE Choline oxidase (EC 1.1.3.17) (Fragment).  
 OS Alcaligenes sp.  
 CC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;  
 CC Alcaligenes.  
 OX NCBI\_TaxID=512;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=81006769; PubMed=6997283;  
 RA Ohta-Fukuyama M., Miyake Y., Eml S., Yamano T.;  
 RT "Identification and properties of the prosthetic group of choline  
 oxidase from Alcaligenes sp."  
 RL J. Biochem. 88:197-203(1980).  
 CC -1- CATALYTIC ACTIVITY: Choline + O(2) -> betaine aldehyde + H(2)O(2).  
 DR PIR: A15398; A15398.  
 KW Oxidoreductase.  
 FT NON\_TER 7  
 SQ SEQUENCE 7 AA; 839 MW; 7A15B1E457644AC0 CRC64;  
 QY Query Match 35.3%; Score 6; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 12 D 12  
 1 D 1

RESULT 14  
 FAR1\_HELTI STANDARD; PRT; 7 AA.  
 ID ID  
 AC P41871;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE FMRFamide-like neuropeptide GDFPLRF-amide.  
 OS Helisoma trivolvis (Snail).  
 CC Eukaryota; Metazoa; Mollusca; Gastropoda; Basommatophora;  
 CC Planorbidae; Helisoma.  
 OX NCBI\_TaxID=27815;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE-Kidney;

RX MEDLINE=94286417; PubMed=7912428;  
 RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;  
 RT "FMRFamide-related peptides from the kidney of the snail, Helisoma  
 trivolvis."  
 RL Peptides 15:31-36(1994).  
 CC -1- FUNCTION: APPEARS TO BE INVOLVED IN OSMOREGULATION BY AFFECTING  
 CC THE KIDNEY, MANTLE AND SKIN.  
 CC -1- TISSUE SPECIFICITY: KIDNEY, SKIN, MANTLE AND THE HEMOLYMPH.  
 CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 CC Neuropeptide; Amidation.  
 FT MOD\_RES 7  
 SQ SEQUENCE 7 AA; 851 MW; 69D40729D76AA810 CRC64;  
 QY Query Match 35.3%; Score 6; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 12 D 12  
 2 D 2

RESULT 15  
 FAR2\_PROCL STANDARD; PRT; 7 AA.  
 ID ID  
 AC P38498;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE Cardioexcitatory FMRFamide homolog DF2.  
 OS Procambarus clarkii (Red swamp crayfish).  
 CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;  
 CC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;  
 CC Astacidea; Astacoidea; Cambaridae; Procambarus.  
 OX NCBI\_TaxID=6728;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE-Pericardial organs;  
 RX MEDLINE=93248032; PubMed=8387183;  
 RA Mercier A.J., Orchard I., Tebruge V., Skerrett M.;  
 RT "Isolation of two FMRFamide-related peptides from crayfish  
 pericardial organs."  
 RL Peptides 14:137-143(1993).  
 CC -1- FUNCTION: INCREASES THE RATE AND AMPLITUDE OF SPONTANEOUS  
 CC CONTRACTIONS OF SEMI-ISOLATED HEARTS. INCREASES THE AMPLITUDE OF  
 CC EXCITATORY POSTSYNAPTIC POTENTIALS IN ABDOMINAL EXTENSOR MUSCLE.  
 CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 CC Neuropeptide; Amidation.  
 FT MOD\_RES 7  
 SQ SEQUENCE 7 AA; 967 MW; 69D40729C4540AC0 CRC64;  
 QY Query Match 35.3%; Score 6; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 12 D 12  
 1 D 1

Search completed: June 5, 2003, 15:56:35  
 Job time : 7.81818 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:32:58 ; Search time 34.9091 Seconds  
(without alignments)  
70.829 Million cell updates/sec

Title: US-09-150-947F-18  
Perfect score: 17  
Sequence: 1 xxxxxxxxxxxd 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SP archaea:\*  
2: SP bacteria:\*  
3: SP fungi:\*  
4: SP human:\*  
5: SP invertebrate:\*  
6: SP mammal:\*  
7: SP mhc:\*  
8: SP organelle:\*  
9: SP phage:\*  
10: SP plant:\*  
11: SP rodent:\*  
12: SP virus:\*  
13: SP vertebrate:\*  
14: SP unclassified:\*  
15: SP virus:\*  
16: SP bacteriophage:\*  
17: SP archaea:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	35.3	5	2	P83073 bacillus ce
2	6	35.3	7	2	O07354 synechococ
3	6	35.3	7	5	P83274 macrobrach
4	6	35.3	7	6	O28742 oryctolagus
5	6	35.3	7	10	O49223 glycine max
6	6	35.3	7	11	O63480 rattus norv
7	6	35.3	7	11	O55184 rattus norv
8	6	35.3	8	2	O92189 neisseria m
9	6	35.3	8	2	O85406 coxiella bu
10	6	35.3	8	2	O9RQ57 buchnera ap
11	6	35.3	8	2	O9RQ49 buchnera ap
12	6	35.3	8	2	O9R772 escherichia
13	6	35.3	8	2	P72279 rhodococcus
14	6	35.3	8	2	O56759 xanthobacte
15	6	35.3	8	2	P83158 anabaena sp
16	6	35.3	8	2	O51594 escherichia

17	6	35.3	8	2	O93SR0	O93SR0 staphylococ
18	6	35.3	8	2	P83152	P83152 anabaena sp
19	6	35.3	8	2	O9R3X0	O9R3X0 planktothri
20	6	35.3	8	2	O53790	O53790 streptococc
21	6	35.3	8	2	O8R5R3	O8R5R3 lactobacilli
22	6	35.3	8	3	O9HDS4	O9HDS4 aspergillus
23	6	35.3	8	3	O9URB9	O9URB9 saccharomyc
24	6	35.3	8	4	O9P285	O9P285 homo sapien
25	6	35.3	8	4	O9UJ50	O9UJ50 homo sapien
26	6	35.3	8	4	O9HAD3	O9HAD3 homo sapien
27	6	35.3	8	4	O15895	O15895 homo sapien
28	6	35.3	8	4	O15900	O15900 homo sapien
29	6	35.3	8	4	O15902	O15902 homo sapien
30	6	35.3	8	4	O9UMC7	O9UMC7 homo sapien
31	6	35.3	8	4	O9UMH9	O9UMH9 homo sapien
32	6	35.3	8	4	O9UDZ4	O9UDZ4 homo sapien
33	6	35.3	8	4	O9P0K3	O9P0K3 homo sapien
34	6	35.3	8	4	O9BYR5	O9BYR5 toxoplasma
35	6	35.3	8	5	O9NMK5	O9NMK5 physarum po
36	6	35.3	8	5	O9UB13	O9UB13 perkinsus a
37	6	35.3	8	5	O94695	O94695 perkinsus a
38	6	35.3	8	5	P83195	P83195 macrobrachi
39	6	35.3	8	5	P83275	P83275 penaeus mon
40	6	35.3	8	5	P83316	P83316 perineleis
41	6	35.3	8	5	O9TWH6	O9TWH6 periplaneta
42	6	35.3	8	5	P82686	P82686 periplaneta
43	6	35.3	8	5	P82687	P82687 periplaneta
44	6	35.3	8	6	O9T778	O9T778 canis famili
45	6	35.3	8	6	O8WNS1	O8WNS1 bos taurus

## ALIGNMENTS

RESULT 1  
P83073 PRELIMINARY; PRT; 5 AA.  
AC P83073;  
DT 01-OCT-2001 (TREMUREL.18, Created)  
DT 01-OCT-2001 (TREMUREL.18, Last sequence update)  
DT 01-OCT-2001 (TREMUREL.18, Last annotation update)  
DE 88 kDa protein (Fragment).  
OS Bacillus cereus.  
OC Bacteria; Firmicutes; Bacillus/clostridium group; Bacillales;  
OC Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1396;  
RN [1]  
RP  
RC STRAIN-NCIMB 11796;  
RA Browne N., Dowds B.C.A.;  
RL Submitted (JUL-2001) to the SWISS-PROT data bank.  
FT NON\_TER 5  
SQ SEQUENCE 5 AA; 623 MW; 6801AAA336F00000 CRC64;

Query Match 35.3%; Score 6; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
Db 3 D 3  
RESULT 2  
O07354 PRELIMINARY; PRT; 7 AA.  
AC O07354;  
DT 01-JUL-1997 (TREMUREL.04, Created)  
DT 01-JUL-1997 (TREMUREL.04, Last sequence update)  
DT 01-DEC-2001 (TREMUREL.19, Last annotation update)  
DE NifK (Fragment).  
GN NifK.  
OS synechococcus sp. (strain PCC 8801 / Rf-1) (Cyanothecae PCC 8801).

OC Bacteria; Cyanobacteria; Chroococcales; Cyanothecae.  
 OX NCBI\_TaxID=41431;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-RF-1;  
 RX MEDLINE=99231861; PubMed=10217509;  
 RA Huang T.C., Lin R.F., Chu M.K., Chen H.M.;  
 RT "Organization and expression of nitrogen-fixation genes in the aerobic  
 RT nitrogen-fixing unicellular cyanobacterium *Synechococcus* sp. strain  
 RF-1."  
 RL Microbiology 145:743-753(1999).  
 DR EMBL: AF003700; AAC35193.1; -.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 7 AA; 849 MW; 7412C72AA9D5B030 CRC64;  
 Query Match 35.3%; Score 6; DB 2; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 12 D 12  
 DB 4 D 4

RESULT 3  
 ID P83274 PRELIMINARY; PRT; 7 AA.  
 AC P83274;  
 DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DE FMFamide-1-like neuropeptide FLP1 (DRNFLR-amide).  
 OS Macrobrachium rosenbergii (Giant fresh water prawn).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Caridea;  
 OC Palaemonoidea; Palaemonidae; Macrobrachium.  
 OX NCBI\_TaxID=79674;  
 RN [1]  
 RP SEQUENCE, AND MASS SPECTROMETRY.  
 RC TISSUE-EYESTALK;  
 RA Sithigorngul P., Sarathongkum W., Jaldeschoey S., Longyant S.,  
 RA Sithigorngul W.;  
 RT "Novel FMFamide-like neuropeptides from the eyestalk of the giant  
 RT freshwater prawn *Macrobrachium rosenbergii*."  
 RL Comp. Biochem. Physiol. 120B:587-595(1998).  
 CC -1- MASS SPECTROMETRY: MW=965.7; METHOD=MALDI.  
 CC -1- SIMILARITY: BELONGS TO THE FARP (FMFAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 CC Neuropeptide: Amidation.  
 KW MOD\_RES 7 7  
 SQ SEQUENCE 7 AA; 967 MW; 69D40729C4540AC0 CRC64;  
 Query Match 35.3%; Score 6; DB 5; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 12 D 12  
 DB 1 D 1

RESULT 4  
 ID Q28742 PRELIMINARY; PRT; 7 AA.  
 AC Q28742;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DE Alpha-myosin heavy chain (Fragment).  
 OS *Oryctolagus cuniculus* (Rabbit).  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84221901; PubMed=6328491;  
 RA Friedman D.J., Umeda P.K., Sinha A.M., Hsu H.-J., Jokovic S.,  
 RA Rabinowitz M.;  
 RT "Characterization of genomic clones specifying rabbit alpha- and beta-  
 RT ventricular myosin heavy chains."  
 RL Proc. Natl. Acad. Sci. U.S.A. 81:3044-3048(1984).  
 DR EMBL: K01698; AAA31415.1; -.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 7 AA; 916 MW; 6B1B1A1E6932680 CRC64;  
 Query Match 35.3%; Score 6; DB 6; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 12 D 12  
 DB 5 D 5

RESULT 5  
 ID O49223 PRELIMINARY; PRT; 7 AA.  
 AC O49223;  
 DT 01-JUN-1998 (TREMBlrel. 06, Created)  
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
 DE HMG-1-like protein (Fragment).  
 OS Glycine max (Soybean).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
 OX NCBI\_TaxID=3847;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. ESSEX; TISSUE=ROOT;  
 RX MEDLINE=91367679; PubMed=1891369;  
 RA Laux T., Goldberg R.B.;  
 RT "A plant DNA binding protein shares highly conserved sequence motifs  
 RT with HMG-box proteins."  
 RL Nucleic Acids Res. 19:4769-4769(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. ESSEX; TISSUE=ROOT;  
 RA Mahalingam R., Knapp H.T.;  
 RL Submitted (Feb-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF047050; AAC03556.1; -.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 7 AA; 850 MW; 6AAAAAB378637810 CRC64;  
 Query Match 35.3%; Score 6; DB 10; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 12 D 12  
 DB 5 D 5

RESULT 6  
 ID O63480 PRELIMINARY; PRT; 7 AA.  
 AC O63480;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DE TR4-NS orphan receptor (Fragment).  
 OS *Rattus norvegicus* (Rat).  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-96198747; PubMed-8612486;  
 RA Yoshikawa T., Makino S., Gao X.M., Xing G.Q., Chuang D.M.,  
 Detera-Wadleigh S.D.;  
 RT "Splice variants of rat TR4 orphan receptor: differential expression  
 of novel sequences in the 5'-untranslated region and C-terminal  
 domain";  
 RT domain";  
 RL Endocrinology 137:1562-1571(1996).  
 DR EMBL: U59125; AAB02827.1; -.  
 KW Receptor.  
 FT NON\_TER  
 SQ SEQUENCE 7 AA; 758 MW; 672AA87864005350 CRC64;  
 QY  
 Db 12 D 12  
 6 D 6  
 Query Match 35.3%; Score 6; DB 11; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 7  
 055184 PRELIMINARY; PRT; 7 AA.  
 AC 055184;  
 DT 01-JUN-1998 (TREMblrel. 06, Created)  
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE Orphan receptor TR4-NS (Fragment).  
 GN TR4.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_Taxid=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SPRAGUE-DAWLEY;  
 RX MEDLINE-96198747; PubMed-8612486;  
 RA Yoshikawa T., Makino S., Gao X.M., Xing G.Q., Chuang D.M.,  
 Detera-Wadleigh S.D.;  
 RT "Splice variants of rat TR4 orphan receptor: differential expression  
 of novel sequences in the 5'-untranslated region and C-terminal  
 domain";  
 RT domain";  
 RL Endocrinology 137:1562-1571(1996).  
 DR EMBL: U59125; AAB02827.1; -.  
 KW Receptor.  
 FT NON\_TER  
 SQ SEQUENCE 7 AA; 663 MW; 6DDA8A787EB05350 CRC64;  
 QY  
 Db 12 D 12  
 6 D 6  
 Query Match 35.3%; Score 6; DB 11; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 8  
 0921E9 PRELIMINARY; PRT; 8 AA.  
 ID 0921E9;  
 AC 0921E9;  
 DT 01-MAY-1999 (TREMblrel. 10, Created)

DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMblrel. 10, Last annotation update)  
 DE Carbamoyl-phosphate synthase subunit B (Fragment).  
 GN CARB.  
 OS Neisseria meningitidis.  
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
 OX NCBI\_Taxid=487;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-1527;  
 RX MEDLINE-95291461; PubMed-7773412;  
 RA Lawson F.S., Billowes F.M., Dillon J.A.;  
 RT "Organization of carbamoyl-phosphate synthase genes in Neisseria  
 gonorrhoeae includes a large, variable intergenic sequence which is  
 RT also present in other Neisseria species";  
 RT Microbiology 141:0-0(0).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-1527;  
 RA Brinkman F.S.L., Francis F.M., Dillon J.R.;  
 RT "Complexity of the variable sequence between the carbamoyl-phosphate  
 RT synthase genes of Neisseria species";  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF029361; AAC78449.1; -.  
 FT NON\_TER  
 SQ SEQUENCE 8 AA; 988 MW; FA372AB1B4032766 CRC64;  
 QY  
 Db 12 D 12  
 6 D 6  
 Query Match 35.3%; Score 6; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 9  
 085406 PRELIMINARY; PRT; 8 AA.  
 ID 085406;  
 AC 085406;  
 DT 01-NOV-1998 (TREMblrel. 08, Created)  
 DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE Hypothetical 1.0 kDa protein (Fragment).  
 OS Coxiella burnetii.  
 OC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;  
 OC Coxiella group; Coxiella.  
 OX NCBI\_Taxid=777;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-NINE MILE PHASE I;  
 RX MEDLINE-98348442; PubMed-9683477;  
 RA Williams H., Jaeger C., Baljer G.;  
 RT "Physical and genetic map of the obligate intracellular bacterium  
 RT Coxiella burnetii";  
 RL J. Bacteriol. 180:3816-3822(1998).  
 DR EMBL: AF064963; AAD09947.1; -.  
 KW Hypothetical protein.  
 FT NON\_TER  
 SQ SEQUENCE 8 AA; 993 MW; 046B5AA453772727 CRC64;  
 QY  
 Db 12 D 12  
 6 D 6  
 Query Match 35.3%; Score 6; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 10  
 09RO57 PRELIMINARY; PRT; 8 AA.  
 ID 09RO57

AC O9R057: 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE Nifs protein homolog (Fragment).  
 GN NIFS.  
 OS Buchnera aphidicola.  
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.  
 OX NCBI\_TaxID=9;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20022990; PubMed=10555290;  
 RA Clark M.A., Moran N.A., Baumann P.;  
 RT "Sequence evolution in bacterial endosymbionts having extreme base  
 RT compositions.";  
 RL MOL. BIOL. EVOL. 16:1586-1598(1999).  
 DR EMBL; AF130812; AAF13797.1; -.  
 FT NON\_TER  
 SQ SEQUENCE 8 AA; 980 MW; F3A73B504771A336 CRC64;

Query Match 35.3%; Score 6; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
 DB 8 D 8

## RESULT 11

O9R049: PRELIMINARY; PRT; 8 AA.  
 AC O9R049: 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
 DE Nifs protein homolog (Fragment).  
 GN NIFS.  
 OS Buchnera aphidicola.  
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.  
 OX NCBI\_TaxID=9;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20022990; PubMed=10555290;  
 RA Clark M.A., Moran N.A., Baumann P.;  
 RT "Sequence evolution in bacterial endosymbionts having extreme base  
 RT compositions.";  
 RL MOL. BIOL. EVOL. 16:1586-1598(1999).  
 DR EMBL; AF130814; AAF13805.1; -.  
 FT NON\_TER  
 SQ SEQUENCE 8 AA; 992 MW; F3A73B5047672336 CRC64;

Query Match 35.3%; Score 6; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
 DB 8 D 8

## RESULT 12

O9R7T2: PRELIMINARY; PRT; 8 AA.  
 AC O9R7T2: 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Hypothetical 1.0 kDa protein (Fragment).  
 GN YOF6.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 Escherichia.

OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE=97061202; PubMed=8905232;  
 RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,  
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,  
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,  
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,  
 RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,  
 RA Yano M., Horinouchi T.;  
 RT "A 718-kb DNA Sequence of Escherichia coli K-12 Genome Corresponding  
 RT to the 12.7-28.0 min Region on the Linkage Map.";  
 RL DNA Res. 3:137-155(1996).  
 DR EMBL; D90705; BAA35310.1; -.  
 FT Hypothetical protein.  
 KW NON\_TER  
 SQ SEQUENCE 8 AA; 964 MW; DF133BDD04B476A CRC64;

Query Match 35.3%; Score 6; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
 DB 1 D 1

## RESULT 13

P72279: PRELIMINARY; PRT; 8 AA.  
 AC P72279: 01-FEB-1997 (TREMBlrel. 02, Created)  
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Biphenyl dioxygenase (Fragment).  
 GN BPHB.  
 OS Rhodococcus globerulus.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacteriaceae; Nocardiaceae; Rhodococcus.  
 OX NCBI\_TaxID=33008;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95255652; PubMed=7737502;  
 RA Asturias J.A., Diaz E., Timmis K.N.;  
 RT "Evolutionary relationship of the biphenyl dioxygenase of the gram-  
 RT positive bacterium Rhodococcus globerulus p6 to multicomponent  
 RT dioxygenases of gram-negative bacteria.";  
 RL Gene 156:11-18(1995).  
 DR EMBL; X80041; CAA56350.1; -.  
 KW Dioxygenase.  
 FT NON\_TER  
 SQ SEQUENCE 8 AA; 989 MW; EBD2CB1AB6D73406 CRC64;

Query Match 35.3%; Score 6; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
 DB 5 D 5

## RESULT 14

O56759: PRELIMINARY; PRT; 8 AA.  
 AC O56759: 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Halocacid dehalogenase (Fragment).  
 GN DHB.  
 OS Xanthobacter autotrophicus.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Hyphomicrobium group; Xanthobacter.

OX NCBI\_TaxID=280;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN-GJ10, AND CV. M50;

RK MEDLINE-95173113; PubMed-7868610;  
 RA Van der Ploeg J., Willemse M., van Hall G., Janssen D.B.;

RT "Adaptation of Xanthobacter autotrophicus GJ10 to bromoacetate due to  
 RT activation and mobilization of the haloacetate dehalogenase gene by  
 RT insertion element IS1247";

RL J. Bacteriol. 177:1348-1356(1995).  
 DR EMBL: X84038; CAA58857.1; -.

FT NON\_TER  
 SQ SEQUENCE 8 AA; 922 MW; F3A9D2D2CDD33056 CRC64;

# Query Match

Best Local Similarity 35.3%; Score 6; DB 2; Length 8;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
 |  
 DB 8 D 8

## RESULT 15

P83158

ID P83158 PRELIMINARY; PRT; 8 AA.

AC P83158;

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Photosystem I iron-sulfur center (Photosystem I subunit VII) (9 kDa  
 DE polypeptide) (PSI-C) (Fragment).

OS Anabaena sp. (strain L33).

OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.

OX NCBI\_TaxID=29412;

RN [1]

RP SEQUENCE.

RA Apte S.K., Uhlemann E., Schmid R., Altendorf K.;

RL Submitted (OCT-2001) to the SWISS-PROT data bank.

CC -!- FUNCTION: APOPROTEIN FOR THE IRON-SULFUR CENTERS FA AND FB OF THE  
 CC PHOTOSYSTEM I COMPLEX.

CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.

CC -!- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF  
 CC 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.

DR InterPro: IPR001450; 4Fe4S\_ferredoxin.

DR PROSITE: PS00198; 4FE4S\_FERREDOXIN; PARTIAL.

KW Photosynthesis; Photosystem I; Iron-sulfur; 4Fe-4S; Metal-binding.

FT NON\_TER  
 SQ SEQUENCE 8 AA; 962 MW; C5B850532D1A1F5 CRC64;

# Query Match

Best Local Similarity 35.3%; Score 6; DB 2; Length 8;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
 |  
 DB 8 D 8

Search completed: June 5, 2003, 15:55:45  
 Job time : 34.9091 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:29:17 ; Search time 123.182 seconds

(without alignments)  
52.340 Million cell updates/sec

Title: US-09-150-947F-13

Perfect score: 31

Sequence: 1 KXXXTXQEXD 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

```

1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	77.4	204	22	US-09-869-136-9
2	24	77.4	209	21	US-09-791-537-103693
3	24	77.4	209	21	US-09-791-537-103695
4	24	77.4	209	21	US-09-791-537-103701
5	24	77.4	209	21	US-09-791-537-103704
6	24	77.4	209	21	US-09-791-537-103705

7	24	77.4	209	21	US-09-791-537-103706	Sequence 103706,
8	24	77.4	209	21	US-09-791-537-103708	Sequence 103708,
9	24	77.4	209	21	US-09-791-537-103712	Sequence 103712,
10	24	77.4	209	21	US-09-791-537-103715	Sequence 103715,
11	24	77.4	209	21	US-09-791-537-103717	Sequence 103717,
12	24	77.4	209	21	US-09-791-537-103719	Sequence 103719,
13	24	77.4	209	21	US-09-791-537-103720	Sequence 103720,
14	24	77.4	209	21	US-09-791-537-103729	Sequence 103729,
15	24	77.4	209	21	US-09-791-537-103733	Sequence 103733,
16	24	77.4	209	21	US-09-791-537-103736	Sequence 103736,
17	24	77.4	209	21	US-09-791-537-103738	Sequence 103738,
18	24	77.4	209	21	US-09-791-537-103740	Sequence 103740,
19	24	77.4	209	21	US-09-791-537-103749	Sequence 103749,
20	24	77.4	209	21	US-09-791-537-120755	Sequence 120755,
21	24	77.4	209	21	US-09-791-537-137418	Sequence 137418,
22	24	77.4	210	21	US-09-791-537-85735	Sequence 85735, A
23	24	77.4	233	22	US-09-869-136-2	Sequence 2, Appl1
24	24	77.4	256	21	US-09-791-537-86827	Sequence 86827, A
25	24	77.4	256	21	US-09-791-537-88724	Sequence 88724, A
26	24	77.4	256	21	US-09-791-537-97492	Sequence 97492, A
27	24	77.4	278	20	US-09-675-784A-11175	Sequence 11175, A
28	24	77.4	603	21	US-09-791-537-25751	Sequence 25751, A
29	24	77.4	603	21	US-09-791-537-41676	Sequence 41676, A
30	24	77.4	604	21	US-09-791-537-12679	Sequence 12679, A
31	24	77.4	604	21	US-09-791-537-12679	Sequence 12679, A
32	24	77.4	955	21	US-09-791-537-100507	Sequence 100507,
33	23	74.2	17	8	US-08-491-746-6	Sequence 6, Appl1
34	23	74.2	17	19	US-09-555-115A-30	Sequence 30, Appl
35	23	74.2	17	19	US-09-555-115A-31	Sequence 31, Appl
36	23	74.2	17	19	US-09-555-115A-32	Sequence 32, Appl
37	23	74.2	17	21	US-09-708-008B-6	Sequence 6, Appl1
38	23	74.2	24	1	PCT-US00-16680A-20	Sequence 20, Appl
39	23	74.2	24	12	US-08-838-413-20	Sequence 20, Appl
40	23	74.2	24	17	US-09-335-581-20	Sequence 20, Appl
41	23	74.2	24	17	US-09-335-581A-20	Sequence 20, Appl
42	23	74.2	24	17	US-09-335-621B-20	Sequence 20, Appl
43	23	74.2	24	17	US-09-335-581-20	Sequence 20, Appl
44	23	74.2	29	1	PCT-US01-00663-32245	Sequence 32245, A
45	23	74.2	29	22	US-09-864-761-43487	Sequence 43487, A

#### ALIGNMENTS

```

RESULT 1
US-09-869-136-9
; Sequence 9, Application US/09869136
; GENERAL INFORMATION:
; APPLICANT: FRASER, JOHN DAVID
; APPLICANT: PROFT, THOMAS
; TITLE OF INVENTION: SUPERANTIGENS
; FILE REFERENCE: 3911-8
; CURRENT APPLICATION NUMBER: US/09/869,136
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: PCT/NZ99/00228
; PRIOR FILING DATE: 1999-12-24
; PRIOR APPLICATION NUMBER: NZ 333589
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; TYPE: PRT
; LENGTH: 204
; ORGANISM: Streptococcus pyogenes
US-09-869-136-9

```

Query Match 77.4%; Score 24; DB 22; Length 204;  
Best local Similarity 50.0%; Pred. No. 5,9e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10  
DB 123 KTTVTAAQED 132

RESULT 2  
US-09-791-537-103693  
; Sequence 103693, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Biomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 103693  
; LENGTH: 209  
; TYPE: PRT  
; ORGANISM: Streptococcus pyogenes  
US-09-791-537-103693

Query Match 77.4%; Score 24; DB 21; Length 209;  
Best Local Similarity 50.0%; Pred. No. 6.1e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10  
| | | |  
DB 125 KTYVTAQIED 134

RESULT 3  
US-09-791-537-103695  
; Sequence 103695, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Biomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 103695  
; LENGTH: 209  
; TYPE: PRT  
; ORGANISM: Streptococcus pyogenes  
US-09-791-537-103695

Query Match 77.4%; Score 24; DB 21; Length 209;  
Best Local Similarity 50.0%; Pred. No. 6.1e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10  
| | | |  
DB 125 KTYVTAQIED 134

RESULT 4  
US-09-791-537-103701  
; Sequence 103701, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Biomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 103701  
; LENGTH: 209  
; TYPE: PRT  
; ORGANISM: Streptococcus pyogenes  
US-09-791-537-103701

Query Match 77.4%; Score 24; DB 21; Length 209;  
Best Local Similarity 50.0%; Pred. No. 6.1e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10  
| | | |  
DB 125 KTYVTAQIED 134

RESULT 5  
US-09-791-537-103704  
; Sequence 103704, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Biomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 103704  
; LENGTH: 209  
; TYPE: PRT  
; ORGANISM: Streptococcus pyogenes  
US-09-791-537-103704

Query Match 77.4%; Score 24; DB 21; Length 209;  
Best Local Similarity 50.0%; Pred. No. 6.1e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10  
| | | |  
DB 125 KTYVTAQIED 134

RESULT 6  
US-09-791-537-103705  
; Sequence 103705, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Biomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 103705  
; LENGTH: 209  
; TYPE: PRT  
; ORGANISM: Streptococcus pyogenes  
US-09-791-537-103705

Query Match 77.4%; Score 24; DB 21; Length 209;  
Best Local Similarity 50.0%; Pred. No. 6.1e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10  
| | | |  
DB 125 KTYVTAQIED 134



```
RESULT 7
US-09-791-537-103706
; Sequence 103706, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 103706
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-09-791-537-103706

Query Match
Best Local Similarity 50.0%; Pred. No. 6.1e+02; Length 209;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10
DB 125 KTTVTAQEXD 134

RESULT 8
US-09-791-537-103708
; Sequence 103708, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 103708
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-09-791-537-103708

Query Match
Best Local Similarity 50.0%; Pred. No. 6.1e+02; Length 209;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10
DB 125 KTTVTAQEXD 134

RESULT 9
US-09-791-537-103712
; Sequence 103712, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 103712
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-09-791-537-103712

Query Match
Best Local Similarity 50.0%; Pred. No. 6.1e+02; Length 209;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10
DB 125 KTTVTAQEXD 134

RESULT 10
US-09-791-537-103715
; Sequence 103715, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 103715
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-09-791-537-103715

Query Match
Best Local Similarity 50.0%; Pred. No. 6.1e+02; Length 209;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10
DB 125 KTTVTAQEXD 134

RESULT 11
US-09-791-537-103717
; Sequence 103717, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 103717
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-09-791-537-103717

Query Match
Best Local Similarity 50.0%; Pred. No. 6.1e+02; Length 209;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10
DB 125 KTTVTAQEXD 134
```

RESULT 12  
US-09-791-537-103719  
; Sequence 103719, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 103719  
; LENGTH: 209  
; TYPE: PRT  
; ORGANISM: Streptococcus pyogenes  
US-09-791-537-103719

Query Match 77.4%; Score 24; DB 21; Length 209;  
Best Local Similarity 50.0%; Pred. No. 6.1e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
| | | | |  
Db 125 KTVYTAQED 134

RESULT 13  
US-09-791-537-103720  
; Sequence 103720, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 103720  
; LENGTH: 209  
; TYPE: PRT  
; ORGANISM: Streptococcus pyogenes  
US-09-791-537-103720

Query Match 77.4%; Score 24; DB 21; Length 209;  
Best Local Similarity 50.0%; Pred. No. 6.1e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
| | | | |  
Db 125 KTVYTAQED 134

RESULT 14  
US-09-791-537-103729  
; Sequence 103729, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 103729

; LENGTH: 209  
; TYPE: PRT  
; ORGANISM: Streptococcus pyogenes  
US-09-791-537-103729

Query Match 77.4%; Score 24; DB 21; Length 209;  
Best Local Similarity 50.0%; Pred. No. 6.1e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
| | | | |  
Db 125 KTVYTAQED 134

RESULT 15  
US-09-791-537-103733  
; Sequence 103733, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 103733  
; LENGTH: 209  
; TYPE: PRT  
; ORGANISM: Streptococcus pyogenes  
US-09-791-537-103733

Query Match 77.4%; Score 24; DB 21; Length 209;  
Best Local Similarity 50.0%; Pred. No. 6.1e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQEXD 10  
| | | | |  
Db 125 KTVYTAQED 134

Search completed: June 5, 2003, 15:47:34  
Job time : 124.182 secs



```
;; CURRENT FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/156,338
;; PRIOR FILING DATE: 1999-09-29
;; NUMBER OF SEQ ID NOS: 13925
;; SEQ ID NO 11175
;; LENGTH: 278
;; TYPE: PRT
;; ORGANISM: Aspergillus fumigatus
US-09-675-784A-11175
```

```
Query Match      77.4%; Score 24; DB 5; Length 278;
Best Local Similarity 50.0%; Pred. No. 85;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 KXXXTXQEXD 10
      | | | | |
DB      126 KSDTDOESD 135
```

## RESULT 3

```
US-10-172-425B-37
;; Sequence 37, Application US/10172425B
;; GENERAL INFORMATION:
;; APPLICANT: Kaempfer, Raymond
;; APPLICANT: Arad, Gila
;; TITLE OF INVENTION: BROAD SPECTRUM ANTAGONISTS AND VACCINES
;; TITLE OF INVENTION: DIRECTED AGAINST PYROGENIC EXOTOXINS
;; FILE REFERENCE: A31967-PCF-USA-A-066031.0164
;; CURRENT APPLICATION NUMBER: US/10/172,425B
;; CURRENT FILING DATE: 2002-06-13
;; PRIOR APPLICATION NUMBER: 09/150,947
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: PCF/1197/00438
;; PRIOR FILING DATE: 1997-12-30
;; PRIOR APPLICATION NUMBER: ISRAEL 119938
;; PRIOR FILING DATE: 1996-12-30
;; NUMBER OF SEQ ID NOS: 57
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 37
;; LENGTH: 14
;; TYPE: PRT
;; ORGANISM: Staphylococcus aureus
US-10-172-425B-37
```

```
Query Match      74.2%; Score 23; DB 6; Length 14;
Best Local Similarity 50.0%; Pred. No. 9.2;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 KXXXTXQEXD 10
      | | | | |
DB      5 KSVTAQELD 14
```

## RESULT 4

```
US-09-134-000C-6190
;; Sequence 6190, Application US/09134000C
;; GENERAL INFORMATION:
;; APPLICANT: Lynn Doucette-Stamm et al
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: 032796-032
;; CURRENT APPLICATION NUMBER: US/09/134,000C
;; CURRENT FILING DATE: 1998-08-13
;; PRIOR APPLICATION NUMBER: US 60/055,778
;; PRIOR FILING DATE: 1997-08-15
;; NUMBER OF SEQ ID NOS: 6812
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 6190
;; LENGTH: 115
;; TYPE: PRT
;; ORGANISM: Enterococcus faecalis
US-09-134-000C-6190
```

```
Query Match      74.2%; Score 23; DB 5; Length 115;
Best Local Similarity 50.0%; Pred. No. 69;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 KXXXTXQEXD 10
      | | | | |
DB      85 KKFSTQEVD 94
```

## RESULT 5

```
US-09-134-000C-6190
;; Sequence 6190, Application US/09134000C
;; GENERAL INFORMATION:
;; APPLICANT: Lynn Doucette-Stamm et al
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: 032796-032
;; CURRENT APPLICATION NUMBER: US/09/134,000C
;; CURRENT FILING DATE: 1998-08-13
;; PRIOR APPLICATION NUMBER: US 60/055,778
;; PRIOR FILING DATE: 1997-08-15
;; NUMBER OF SEQ ID NOS: 6812
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 6190
;; LENGTH: 115
;; TYPE: PRT
;; ORGANISM: Enterococcus faecalis
US-09-134-000C-6190
```

```
Query Match      74.2%; Score 23; DB 5; Length 115;
Best Local Similarity 50.0%; Pred. No. 69;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 KXXXTXQEXD 10
      | | | | |
DB      85 KKFSTQEVD 94
```

## RESULT 6

```
US-10-424-599-184221
;; Sequence 184221, Application US/10424599
;; GENERAL INFORMATION:
;; APPLICANT: La Rosa Thomas J
;; APPLICANT: Kovalic David K
;; APPLICANT: Zhou Yihua
;; APPLICANT: Cao Yongwei
;; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
;; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
;; FILE REFERENCE: 38-21(33223)B
;; CURRENT APPLICATION NUMBER: US/10/424,599
;; CURRENT FILING DATE: 2003-04-28
;; NUMBER OF SEQ ID NOS: 285684
;; SEQ ID NO 184221
;; LENGTH: 134
;; TYPE: PRT
;; ORGANISM: Glycine max
;; FEATURE:
;; OTHER INFORMATION: Clone ID: PAT_MRT3847_137368C.1.pep
US-10-424-599-184221
```

```
Query Match      74.2%; Score 23; DB 6; Length 134;
Best Local Similarity 50.0%; Pred. No. 80;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 KXXXTXQEXD 10
      | | | | |
DB      21 KTHDTSQELD 30
```

## RESULT 7

```
US-10-203-536-1
;; Sequence 1, Application US/10203536
;; GENERAL INFORMATION:
```

APPLICANT: LEE, Hong-Kyun  
APPLICANT: PARK, Yong-Ho  
APPLICANT: HAN, Kyu-Boem  
APPLICANT: CHANG, Byoung-Sun  
APPLICANT: LEE, Yong-Jun  
TITLE OF INVENTION: Staphylococcal Enterotoxin SEC-SER, Expression Vector and Host Cell  
FILE REFERENCE: 0808-0329P  
CURRENT FILING DATE: 2002-08-12  
CURRENT APPLICATION NUMBER: US/10/203,536  
PRIOR FILING DATE: 2000-10-31  
PRIOR APPLICATION NUMBER: PCT/KR00/01241  
PRIOR FILING DATE: 2000-02-17  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 1  
LENGTH: 228  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Staphylococcal enterotoxin SEC-SER  
US-10-203-536-1

Query Match 74.2%; Score 23; DB 6; Length 228;  
Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQXND 10  
| | | | |  
Db 140 KKSXTAQELD 149

RESULT 8  
US-09-751-708A-12  
Sequence 12, Application US/09/51708A  
GENERAL INFORMATION:  
APPLICANT: TERMAN, David S  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE  
FILE REFERENCE: 751708  
CURRENT APPLICATION NUMBER: US/09/751,708A  
CURRENT FILING DATE: 2002-10-15  
PRIOR APPLICATION NUMBER: US 60/173,371  
PRIOR FILING DATE: 1999-12-28  
NUMBER OF SEQ ID NOS: 166  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 12  
LENGTH: 266  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-751-708A-12

Query Match 74.2%; Score 23; DB 5; Length 266;  
Best Local Similarity 50.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQXND 10  
| | | | |  
Db 178 KKSXTAQELD 187

RESULT 9  
US-10-424-599-179861  
Sequence 179861, Application US/10424599  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599

CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 179861  
LENGTH: 281  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_13342C.1.pap  
US-10-424-599-179861

Query Match 74.2%; Score 23; DB 6; Length 281;  
Best Local Similarity 50.0%; Pred. No. 1.6e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQXND 10  
| | | | |  
Db 176 KKSXTAQELD 185

RESULT 10  
US-60-440-068-124  
Sequence 124, Application US/60440068  
GENERAL INFORMATION:  
APPLICANT: MADLER, STEVEN G.  
APPLICANT: CARMAN, JULIE  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE  
FILE REFERENCE: 3053-4191  
CURRENT APPLICATION NUMBER: US/60/440,068  
CURRENT FILING DATE: 2003-01-14  
NUMBER OF SEQ ID NOS: 746  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 124  
LENGTH: 324  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-60-440-068-124

Query Match 74.2%; Score 23; DB 7; Length 324;  
Best Local Similarity 50.0%; Pred. No. 1.8e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KXXXTXQXND 10  
| | | | |  
Db 193 KKSXTAQELD 202

RESULT 11  
US-10-431-652-7705  
Sequence 7705, Application US/10431652  
GENERAL INFORMATION:  
APPLICANT: Breton, Gary L.  
APPLICANT: Bush, David  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: PAT03-08  
CURRENT APPLICATION NUMBER: US/10/431,652  
CURRENT FILING DATE: 2003-05-06  
PRIOR APPLICATION NUMBER: US 09/328,352  
PRIOR FILING DATE: 1999-06-04  
PRIOR APPLICATION NUMBER: US 60/088,701  
PRIOR FILING DATE: 1998-06-09  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 7705  
LENGTH: 627  
TYPE: PRT  
ORGANISM: Acinetobacter baumannii  
US-10-431-652-7705

Query Match 74.2%; Score 23; DB 6; Length 627;  
Best Local Similarity 50.0%; Pred. No. 3.5e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10  
DB 355 KSKKTAQOELD 364

RESULT 12  
US-09-150-947F-3

```

; Sequence 3, Application US/09150947F
; GENERAL INFORMATION:
; APPLICANT: Kaempfer, Raymond
; APPLICANT: Arad, Gila
; TITLE OF INVENTION: BROAD SPECTRUM PYROGENIC EXOTOXINS ANTAGONISTS VACCINES
; FILE REFERENCE: A31967-PCT-USA-A 066031.0102
; CURRENT APPLICATION NUMBER: US/09/150,947F
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: PCT/IL97/00438
; PRIOR FILING DATE: 1997-12-30
; PRIOR APPLICATION NUMBER: ISRAEL 119938
; PRIOR FILING DATE: 1996-12-30
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-150-947F-3

```

Query Match 71.0%; Score 22; DB 5; Length 10;  
Best Local Similarity 50.0%; Pred. No. 13;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10  
DB 1 KKKATVQOELD 10

RESULT 13  
US-09-150-947F-4

```

; Sequence 4, Application US/09150947F
; GENERAL INFORMATION:
; APPLICANT: Kaempfer, Raymond
; APPLICANT: Arad, Gila
; TITLE OF INVENTION: BROAD SPECTRUM PYROGENIC EXOTOXINS ANTAGONISTS VACCINES
; FILE REFERENCE: A31967-PCT-USA-A 066031.0102
; CURRENT APPLICATION NUMBER: US/09/150,947F
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: PCT/IL97/00438
; PRIOR FILING DATE: 1997-12-30
; PRIOR APPLICATION NUMBER: ISRAEL 119938
; PRIOR FILING DATE: 1996-12-30
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-150-947F-4

```

Query Match 71.0%; Score 22; DB 5; Length 10;  
Best Local Similarity 50.0%; Pred. No. 13;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10  
DB 1 KKKATVQOELD 10

RESULT 14  
US-10-172-425B-3  
; Sequence 3, Application US/10172425B  
; GENERAL INFORMATION:  
; APPLICANT: Kaempfer, Raymond  
; APPLICANT: Arad, Gila

```

; TITLE OF INVENTION: BROAD SPECTRUM ANTAGONISTS AND VACCINES
; FILE REFERENCE: A31967-PCT-USA-A 066031.0164
; CURRENT APPLICATION NUMBER: US/10/172,425B
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 09/150,947
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: PCT/IL97/00438
; PRIOR FILING DATE: 1997-12-30
; PRIOR APPLICATION NUMBER: ISRAEL 119938
; PRIOR FILING DATE: 1996-12-30
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-172-425B-3

```

Query Match 71.0%; Score 22; DB 6; Length 10;  
Best Local Similarity 50.0%; Pred. No. 13;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10  
DB 1 KKKATVQOELD 10

```

RESULT 15
US-10-172-425B-4
; Sequence 4, Application US/10172425B
; GENERAL INFORMATION:
; APPLICANT: Kaempfer, Raymond
; APPLICANT: Arad, Gila
; TITLE OF INVENTION: BROAD SPECTRUM ANTAGONISTS AND VACCINES
; FILE REFERENCE: A31967-PCT-USA-A 066031.0164
; CURRENT APPLICATION NUMBER: US/10/172,425B
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 09/150,947
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: PCT/IL97/00438
; PRIOR FILING DATE: 1997-12-30
; PRIOR APPLICATION NUMBER: ISRAEL 119938
; PRIOR FILING DATE: 1996-12-30
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-172-425B-4

```

Query Match 71.0%; Score 22; DB 6; Length 10;  
Best Local Similarity 50.0%; Pred. No. 13;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXXXTXQEXD 10  
DB 1 KKKATVQOELD 10

Search completed: June 5, 2003, 15:52:21  
Job time : 32.4848 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:29:17 ; Search time 123.182 seconds

(without alignments)  
52.340 Million cell updates/sec

Title: US-09-150-947f-14

Perfect score: 26

Sequence: 1 KXXXXXXD 10

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US09\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US10\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US11\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US12\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US13\_COMB.pep.\*  
9: /cgn2\_6/ptodata/1/paa/US14\_COMB.pep.\*  
10: /cgn2\_6/ptodata/1/paa/US15\_COMB.pep.\*  
11: /cgn2\_6/ptodata/1/paa/US16\_COMB.pep.\*  
12: /cgn2\_6/ptodata/1/paa/US17\_COMB.pep.\*  
13: /cgn2\_6/ptodata/1/paa/US18\_COMB.pep.\*  
14: /cgn2\_6/ptodata/1/paa/US19\_COMB.pep.\*  
15: /cgn2\_6/ptodata/1/paa/US20\_COMB.pep.\*  
16: /cgn2\_6/ptodata/1/paa/US21\_COMB.pep.\*  
17: /cgn2\_6/ptodata/1/paa/US22\_COMB.pep.\*  
18: /cgn2\_6/ptodata/1/paa/US23\_COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/US24\_COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US25\_COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US26\_COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US27\_COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US28\_COMB.pep.\*  
24: /cgn2\_6/ptodata/1/paa/US29\_COMB.pep.\*  
25: /cgn2\_6/ptodata/1/paa/US30\_COMB.pep.\*  
26: /cgn2\_6/ptodata/1/paa/US31\_COMB.pep.\*  
27: /cgn2\_6/ptodata/1/paa/US32\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	73.1	310	27	US-60-360-039-12686
2	19	73.1	358	21	US-09-733-089-16924
3	19	73.1	358	22	US-09-816-660-16924
4	19	73.1	378	19	US-09-513-996A-43459
5	19	73.1	378	21	US-09-708-427-18784
6	19	73.1	378	21	US-09-791-537-114720

7	19	73.1	378	27	US-60-356-051-2815	Sequence 2815, Ap
8	18	69.2	35	26	US-10-221-279-10068	Sequence 10068, A
9	18	69.2	43	1	PCT-US99-22853B-42	Sequence 42, Appl
10	18	69.2	46	1	PCT-US00-07527-111	Sequence 111, Appl
11	18	69.2	46	1	PCT-US00-07527-112	Sequence 112, Appl
12	18	69.2	54	1	PCT-US99-22853B-41	Sequence 41, Appl
13	18	69.2	79	1	PCT-US99-22853B-40	Sequence 40, Appl
14	18	69.2	87	15	US-09-134-000-5378	Sequence 5378, Ap
15	18	69.2	89	21	US-09-758-472-8508	Sequence 8508, Ap
16	18	69.2	89	26	US-10-235-926-8508	Sequence 8508, Ap
17	18	69.2	100	19	US-09-565-177A-13	Sequence 13, Appl
18	18	69.2	100	21	US-09-732-210-1170	Sequence 1170, Ap
19	18	69.2	100	21	US-09-791-537-23478	Sequence 23478, A
20	18	69.2	100	27	US-60-169-340-1170	Sequence 1170, Ap
21	18	69.2	122	20	US-09-617-681A-4242	Sequence 4242, Ap
22	18	69.2	131	21	US-09-708-427-46061	Sequence 46061, A
23	18	69.2	140	16	US-09-748-766-25982	Sequence 25982, A
24	18	69.2	140	27	US-60-096-409-25982	Sequence 25982, A
25	18	69.2	141	20	US-09-617-681A-4241	Sequence 4241, Ap
26	18	69.2	152	19	US-09-565-177A-4	Sequence 4, Appl
27	18	69.2	152	19	US-09-565-177A-12	Sequence 12, Appl
28	18	69.2	152	19	US-09-565-177A-12	Sequence 12, Appl
29	18	69.2	160	21	US-09-708-427-20661	Sequence 20661, A
30	18	69.2	171	21	US-09-708-427-20661	Sequence 20660, A
31	18	69.2	173	1	PCT-US01-08631-50628	Sequence 50628, A
32	18	69.2	200	22	US-09-809-391-495	Sequence 495, App
33	18	69.2	200	22	US-09-882-171-495	Sequence 495, App
34	18	69.2	200	25	US-10-164-861-495	Sequence 495, App
35	18	69.2	205	27	US-60-389-987-655	Sequence 655, App
36	18	69.2	205	27	US-60-412-418-655	Sequence 655, App
37	18	69.2	206	27	US-60-389-987-2179	Sequence 2179, Ap
38	18	69.2	206	27	US-60-412-418-2179	Sequence 2179, Ap
39	18	69.2	237	1	PCT-US01-14826-218	Sequence 218, Appl
40	18	69.2	237	13	US-08-933-750-32	Sequence 32, Appl
41	18	69.2	237	13	US-08-933-750-32	Sequence 32, Appl
42	18	69.2	237	22	US-09-818-865-32	Sequence 32, Appl
43	18	69.2	237	22	US-09-840-787-32	Sequence 32, Appl
44	18	69.2	237	26	US-10-205-823-91	Sequence 91, Appl
45	18	69.2	247	8	US-08-467-152-18	Sequence 18, Appl

## ALIGNMENTS

RESULT 1  
US-60-360-039-12686  
Sequence 12686, Application US/60360039  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Chen, Xianfeng  
APPLICANT: Goldman, Barry S.  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)A  
CURRENT APPLICATION NUMBER: US/60/360, 039  
CURRENT FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 12686  
LENGTH: 310  
TYPE: PRT  
ORGANISM: Aspergillus nidulans  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(310)  
OTHER INFORMATION: unsure at all Xaa locations  
US-60-360-039-12686

Query Match 73.1% Score 19; DB 27; Length 310;  
Best Local Similarity 40.0%; Pred. No. 2.4e+03;  
Matches 4; Conservative 0; Mismatches 6; Indels 0;

QY 1 KKKXXXXXLD 10  
DB 73 KKTAAAFALD 82

## RESULT 2

US-09-733-089-16924  
Sequence 16924, Application US/09733089  
GENERAL INFORMATION:  
APPLICANT: Dotson, Stanton B.  
APPLICANT: Kovalic, David K.  
APPLICANT: Liu, Jingdong  
APPLICANT: Lutfiyya, Linda L.  
APPLICANT: McIninch, James  
APPLICANT: Wu, Wei  
TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With  
FILE REFERENCE: 38-21(15300)D  
CURRENT APPLICATION NUMBER: US/09/733,089  
CURRENT FILING DATE: 2000-12-11  
PRIOR APPLICATION NUMBER: US 09/474,435  
PRIOR FILING DATE: 1999-12-28  
PRIOR APPLICATION NUMBER: US 09/654,617  
PRIOR FILING DATE: 2000-09-05  
PRIOR APPLICATION NUMBER: US 09/620,392  
PRIOR FILING DATE: 2000-07-19  
NUMBER OF SEQ ID NOS: 24143  
SEQ ID NO 16924  
LENGTH: 358  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-09-733-089-16924

Query Match 73.1%; Score 19; DB 21; Length 358;  
Best Local Similarity 40.0%; Pred. No. 2.8e+03;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKKXXXXXLD 10  
DB 42 KKAADAAALD 51

## RESULT 3

US-09-816-660-16924  
Sequence 16924, Application US/09816660  
GENERAL INFORMATION:  
APPLICANT: Dotson, Stanton B.  
APPLICANT: Kovalic, David K.  
APPLICANT: Liu, Jingdong  
APPLICANT: Lutfiyya, Linda L.  
APPLICANT: McIninch, James  
APPLICANT: Wu, Wei  
TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With  
FILE REFERENCE: 38-21(15300)D  
CURRENT APPLICATION NUMBER: US/09/816,660  
CURRENT FILING DATE: 2001-03-26  
PRIOR APPLICATION NUMBER: US 09/474,435  
PRIOR FILING DATE: 1999-12-28  
PRIOR APPLICATION NUMBER: US 09/654,617  
PRIOR FILING DATE: 2000-09-05  
PRIOR APPLICATION NUMBER: US 09/733,089  
PRIOR FILING DATE: 2000-12-11  
PRIOR APPLICATION NUMBER: ) US 09/684,016  
PRIOR FILING DATE: 2000-10-10  
PRIOR APPLICATION NUMBER: US 09/620,392  
PRIOR FILING DATE: 2000-07-19  
NUMBER OF SEQ ID NOS: 24143  
SEQ ID NO 16924  
LENGTH: 358  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-09-816-660-16924

QY 1 KKKXXXXXLD 10  
DB 42 KKAADAAALD 51

Query Match 73.1%; Score 19; DB 22; Length 358;  
Best Local Similarity 40.0%; Pred. No. 2.8e+03;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKKXXXXXLD 10  
DB 42 KKAADAAALD 51

## RESULT 4

US-09-513-996A-43459  
Sequence 43459, Application US/09513996A  
GENERAL INFORMATION:  
APPLICANT: N. ALEXANDROV et al.  
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
FILE REFERENCE: 2750-709P  
CURRENT APPLICATION NUMBER: US/09/513,996A  
CURRENT FILING DATE: 2000-02-25  
NUMBER OF SEQ ID NOS: 81028  
SEQ ID NO 43459  
LENGTH: 378  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: 1..378  
OTHER INFORMATION: any n or Xaa - unknown  
FEATURE:  
OTHER INFORMATION: Location 1..378 / Ceres Seq. ID 1985639  
US-09-513-996A-43459

Query Match 73.1%; Score 19; DB 19; Length 378;  
Best Local Similarity 40.0%; Pred. No. 3e+03;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKKXXXXXLD 10  
DB 13 KKAADAAALD 22

## RESULT 5

US-09-708-427-18784  
Sequence 18784, Application US/09708427  
GENERAL INFORMATION:  
APPLICANT: N. ALEXANDROV et al.  
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
FILE REFERENCE: 2750-1243P  
CURRENT APPLICATION NUMBER: US/09/708,427  
CURRENT FILING DATE: 2000-11-09  
NUMBER OF SEQ ID NOS: 85364  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 18784  
LENGTH: 378  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: misc. feature  
LOCATION: 1..378  
OTHER INFORMATION: Xaa is any amino acid  
NAME/KEY: misc. feature  
LOCATION: 1..378  
OTHER INFORMATION: Ceres Seq. ID 1834815  
US-09-708-427-18784

Query Match 73.1%; Score 19; DB 21; Length 378;  
Best Local Similarity 40.0%; Pred. No. 3e+03;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKKXXXXXLD 10  
DB 13 KKAADAAALD 22



Db 13 KKAADAAALD 22

## RESULT 6

US-09-791-537-114720  
; Sequence 114720, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: Patent version 3.0  
; SEQ ID NO 114720  
; LENGTH: 378  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-791-537-114720

Query Match 73.1%; Score 19; DB 21; Length 378;  
Best Local Similarity 40.0%; Pred. No. 3e+03;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKKXXXXXLD 10  
Db 13 KKAADAAALD 22

## RESULT 7

US-60-356-051-2815  
; Sequence 2815, Application US/60356051  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Technology LLC  
; APPLICANT: Luffiyya, Linda L  
; TITLE OF INVENTION: NUCLEIC ACIDS AND OTHER MOLECULES ASSOCIATED WITH  
; FILE REFERENCE: 38-21 (15300)I  
; CURRENT APPLICATION NUMBER: US/60/356,051  
; CURRENT FILING DATE: 2002-02-11  
; NUMBER OF SEQ ID NOS: 2926  
; SOFTWARE: Patent version 3.1  
; SEQ ID NO 2815  
; LENGTH: 378  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-60-356-051-2815

Query Match 73.1%; Score 19; DB 27; Length 378;  
Best Local Similarity 40.0%; Pred. No. 3e+03;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKKXXXXXLD 10  
Db 13 KKAADAAALD 22

## RESULT 8

US-10-221-279-10068  
; Sequence 10068, Application US/10221279  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides  
; FILE REFERENCE: 21272-046  
; CURRENT APPLICATION NUMBER: US/10/221,279  
; CURRENT FILING DATE: 2002-09-06  
; PRIOR APPLICATION NUMBER: 09/574,454  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: 09/519,705  
; PRIOR FILING DATE: 2000-03-07

NUMBER OF SEQ ID NOS: 12360  
; SOFTWARE: Custom  
; SEQ ID NO 10068  
; LENGTH: 35  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-221-279-10068

Query Match 69.2%; Score 18; DB 26; Length 35;  
Best Local Similarity 40.0%; Pred. No. 4.5e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKKXXXXXLD 10  
Db 3 KKYSSQTSLD 12

## RESULT 9

PCT-US99-22853B-42  
; Sequence 42, Application PC/TUS9922853B  
; GENERAL INFORMATION:  
; APPLICANT: Ceres, Inc.  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding  
; FILE REFERENCE: 2750-0569f(PC)  
; CURRENT APPLICATION NUMBER: PCT/US99/22853B  
; CURRENT FILING DATE: 1999-10-05  
; NUMBER OF SEQ ID NOS: 3938  
; SOFTWARE: MS word 97  
; SEQ ID NO 42  
; LENGTH: 43  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; OTHER INFORMATION: LOCATION 1..43, Ceres Seq. ID 1685642  
; LOCATION: (1)..(43)  
; OTHER INFORMATION: any Xaa - any amino acid, unknown or other  
PCT-US99-22853B-42

Query Match 69.2%; Score 18; DB 1; Length 43;  
Best Local Similarity 40.0%; Pred. No. 5.7e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKKXXXXXLD 10  
Db 5 KKLTTDALD 14

## RESULT 10

PCT-US00-07527-111  
; Sequence 111, Application PC/TUS0007527  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: 47 Human Secreted Proteins  
; FILE REFERENCE: P5529PCT  
; CURRENT APPLICATION NUMBER: PCT/US00/07527  
; CURRENT FILING DATE: 2000-03-22  
; EARLIER APPLICATION NUMBER: 60/126,600  
; EARLIER FILING DATE: 1999-03-26  
; EARLIER APPLICATION NUMBER: 60/171,550  
; EARLIER FILING DATE: 1999-11,550  
; NUMBER OF SEQ ID NOS: 138  
; SOFTWARE: Patent Ver. 2.0  
; SEQ ID NO 111  
; LENGTH: 46  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US00-07527-111

Query Match 69.2%; Score 18; DB 1; Length 46;  
Best Local Similarity 40.0%; Pred. No. 6.1e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXXXXXXXXD 10  
DB 19 KKSAAESMLD 28

## RESULT 11

PCT-US00-07527-112  
; Sequence 112, Application PC/TUS0007527  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: 47 Human Secreted Proteins  
; FILE REFERENCE: PS529PCT  
; CURRENT APPLICATION NUMBER: PCT/US00/07527  
; CURRENT FILING DATE: 2000-03-22  
; EARLIER APPLICATION NUMBER: 60/126,600  
; EARLIER FILING DATE: 1999-03-26  
; EARLIER APPLICATION NUMBER: 60/171,550  
; EARLIER FILING DATE: 1999-171,550  
; NUMBER OF SEQ ID NOS: 138  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 112  
; LENGTH: 46  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US00-07527-112

Query Match 69.2%; Score 18; DB 1; Length 46;  
Best Local Similarity 40.0%; Pred. No. 6.1e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXXXXXXXXD 10  
DB 19 KKSAAESMLD 28

## RESULT 12

PCT-US99-22853B-41  
; Sequence 41, Application PC/TUS9922853B  
; GENERAL INFORMATION:  
; APPLICANT: Ceres, Inc.  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding  
; FILE REFERENCE: 2750-0569F(PC)  
; CURRENT APPLICATION NUMBER: PCT/US99/22853B  
; CURRENT FILING DATE: 1999-10-05  
; NUMBER OF SEQ ID NOS: 3938  
; SOFTWARE: MS Word 97  
; SEQ ID NO 41  
; LENGTH: 54  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; OTHER INFORMATION: LOCATION 1..54, Ceres Seq. ID 1685641  
PCT-US99-22853B-41

Query Match 69.2%; Score 18; DB 1; Length 54;  
Best Local Similarity 40.0%; Pred. No. 7.3e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXXXXXXXXD 10  
DB 16 KKLTTTDLAD 25

RESULT 13  
PCT-US99-22853B-40  
; Sequence 40, Application PC/TUS9922853B  
; GENERAL INFORMATION:  
; APPLICANT: Ceres, Inc.

; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding  
; FILE REFERENCE: 2750-0569F(PC)  
; CURRENT APPLICATION NUMBER: PCT/US99/22853B  
; CURRENT FILING DATE: 1999-10-05  
; NUMBER OF SEQ ID NOS: 3938  
; SOFTWARE: MS Word 97  
; SEQ ID NO 40  
; LENGTH: 79  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; OTHER INFORMATION: LOCATION 1..79, Ceres Seq. ID 1685640  
; NAME/KEY: UNSURE  
; LOCATION: (1)..(79)  
; OTHER INFORMATION: any Xaa - any amino acid, unknown or other  
PCT-US99-22853B-40

Query Match 69.2%; Score 18; DB 1; Length 79;  
Best Local Similarity 40.0%; Pred. No. 1.1e+03;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXXXXXXXXD 10  
DB 41 KKLTTTDLAD 50

## RESULT 14

US-09-134-000-5378  
; Sequence 5378, Application US/09134000A  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS  
; FILE REFERENCE: GTC-005  
; CURRENT APPLICATION NUMBER: US/09/134,000A  
; CURRENT FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 6810  
; SEQ ID NO 5378  
; LENGTH: 87  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-134-000-5378

Query Match 69.2%; Score 18; DB 15; Length 87;  
Best Local Similarity 40.0%; Pred. No. 1.2e+03;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXXXXXXXXD 10  
DB 44 KKSRAFAASLD 53

## RESULT 15

US-09-758-472-8508  
; Sequence 8508, Application US/09758472  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PH001  
; CURRENT APPLICATION NUMBER: US/09/758,472  
; CURRENT FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; NUMBER OF SEQ ID NOS: 9632  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 8508  
; LENGTH: 89  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:

Fri Jun 6 09:14:34 2003

```

; NAME/KEY: SITE
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (4)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (18)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-758-472-8508
    
```

```

Query Match      69.2%; Score 18; DB 21; Length 89;
Best Local Similarity 40.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
    
```

```

QY      1 KKKXXXXXD 10
         ||  ||
Db      43 KKAISALD 52
    
```

Search completed: June 5, 2003, 15:47:36  
 Job time : 125.182 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: June 5, 2003, 15:31:47 ; Search time 28.4848 Seconds  
(without alignments)  
74.544 Million cell updates/sec

Title: US-09-150-947f-14  
Perfect score: 26  
Sequence: 1 KXXXXXXXXX 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1012620 seqs, 212337834 residues

Total number of hits satisfying chosen parameters: 1012620

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database : Pending\_Patents\_AA.New.\*  
1: /cgn2\_6/pdata/1/paa/PC7\_NEW\_COMB.pep.\*  
2: /cgn2\_6/pdata/1/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/pdata/1/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/pdata/1/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/pdata/1/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/pdata/1/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/pdata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	73.1	310	6	US-10-369-493-12686 Sequence 12686, A
2	19	73.1	378	6	US-10-361-942-986 Sequence 986, App
3	18	69.2	87	5	US-09-134-000C-5378 Sequence 5378, Ap
4	18	69.2	87	5	US-09-134-000C-5378 Sequence 5378, Ap
5	18	69.2	100	6	US-10-421-684-1170 Sequence 1170, Ap
6	18	69.2	109	5	US-09-724-676-55061 Sequence 55061, A
7	18	69.2	109	5	US-09-724-676-55061 Sequence 55061, A
8	18	69.2	147	5	US-09-513-999C-6288 Sequence 6288, Ap
9	18	69.2	147	5	US-09-513-999C-6288 Sequence 6288, Ap
10	18	69.2	152	7	US-09-949-016-6460 Sequence 6460, Ap
11	18	69.2	152	7	US-09-949-016-6460 Sequence 6460, Ap
12	18	69.2	161	5	US-09-724-676-55060 Sequence 55060, A
13	18	69.2	161	5	US-09-724-676-55060 Sequence 55060, A
14	18	69.2	252	6	US-10-282-122A-52371 Sequence 52371, A
15	18	69.2	254	6	US-10-431-652-7651 Sequence 7651, A
16	18	69.2	315	6	US-10-431-652-4290 Sequence 4290, Ap
17	18	69.2	318	6	US-10-272-490-92 Sequence 90, Appl
18	18	69.2	318	6	US-10-272-490-92 Sequence 90, Appl
19	18	69.2	328	6	US-10-363-616-281 Sequence 281, Appl
20	18	69.2	341	6	US-10-272-490-68 Sequence 68, Appl
21	18	69.2	341	6	US-10-272-490-68 Sequence 68, Appl
22	18	69.2	341	6	US-10-214-473-68 Sequence 68, Appl
23	18	69.2	341	6	US-10-214-473-68 Sequence 68, Appl
24	18	69.2	406	6	US-10-425-114-70484 Sequence 70484, A
25	18	69.2	768	6	US-10-218-140-4424 Sequence 4424, A
26	18	69.2	837	6	US-10-369-493-22296 Sequence 22296, A

27	17	65.4	12	6	US-10-172-425B-53	Sequence 53, Appl
28	17	65.4	14	6	US-10-172-425B-37	Sequence 37, Appl
29	17	65.4	14	6	US-10-172-425B-49	Sequence 49, Appl
30	17	65.4	53	6	US-10-424-599-271937	Sequence 271937, Ap
31	17	65.4	57	6	US-10-424-599-172320	Sequence 172320, Ap
32	17	65.4	70	6	US-10-424-599-233335	Sequence 233335, Ap
33	17	65.4	77	6	US-10-424-599-225608	Sequence 225608, Ap
34	17	65.4	85	6	US-10-424-599-194596	Sequence 194596, Ap
35	17	65.4	106	6	US-09-664-408A-3042	Sequence 3042, Ap
36	17	65.4	109	6	US-10-424-599-167306	Sequence 167306, Ap
37	17	65.4	128	6	US-10-357-886-44	Sequence 44, Appl
38	17	65.4	142	5	US-09-857-826B-18	Sequence 18, Appl
39	17	65.4	143	6	US-10-276-774-2588	Sequence 2588, Ap
40	17	65.4	152	6	US-10-417-884-3744	Sequence 3744, Ap
41	17	65.4	160	6	US-10-424-599-170060	Sequence 170060, Ap
42	17	65.4	178	6	US-10-264-237-2676	Sequence 2676, Ap
43	17	65.4	170	6	US-10-282-122A-44631	Sequence 44631, A
44	17	65.4	189	6	US-10-094-149-2347	Sequence 2347, Ap
45	17	65.4	192	5	US-09-675-784A-8051	Sequence 8051, Ap

## ALIGNMENTS

RESULT 1  
US-10-369-493-12686  
Sequence 12686, Application US/10369493  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 4/374  
SEQ ID NO 12686  
LENGTH: 310  
TYPE: PRT  
ORGANISM: Aspergillus nidulans  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(310)  
OTHER INFORMATION: unsure at all Xaa locations  
US-10-369-493-12686

Query Match 73.18: Score 19; DB 6; Length 310;  
Best Local Similarity 40.08; Pred. No. 1.7e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXXXXXXXXX 10  
DB 73 KTTAFAAD 82

RESULT 2  
US-10-361-942-986  
Sequence 986, Application US/10361942  
GENERAL INFORMATION:  
APPLICANT: Lutfiyya, Linda L.  
TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With  
FILE REFERENCE: 38-21(15300)K  
CURRENT APPLICATION NUMBER: US/10/361,942  
CURRENT FILING DATE: 2003-02-10  
PRIOR APPLICATION NUMBER: US 60/356,051  
PRIOR FILING DATE: 2002-02-11  
NUMBER OF SEQ ID NOS: 2906

```
;; SEQ ID NO 986
;; LENGTH: 378
;; TYPE: PRT
;; ORGANISM: Arabidopsis thaliana
;; FEATURE:
;;   OTHER INFORMATION: Clone ID: AT_T6A23.C1.p24.tg
US-10-361-942-986
```

```
Query Match          73.1%; Score 19; DB 6; Length 378;
Best Local Similarity 40.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
OY 1 KKKXXXXXLD 10
    ||
DB 13 KKAADAAALD 22
```

```
RESULT 3
US-09-134-000C-5378
; Sequence 5378, Application US/09134000C
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucelte-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5378
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5378
```

```
Query Match          69.2%; Score 18; DB 5; Length 87;
Best Local Similarity 40.0%; Pred. No. 1e+02;
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
OY 1 KKKXXXXXLD 10
    ||
DB 44 KKSRAAFASLD 53
```

```
RESULT 4
US-09-134-000C-5378
; Sequence 5378, Application US/09134000C
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucelte-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5378
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5378
```

```
Query Match          69.2%; Score 18; DB 5; Length 87;
Best Local Similarity 40.0%; Pred. No. 1e+02;
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
OY 1 KKKXXXXXLD 10
    ||
DB 44 KKSRAAFASLD 53
```

```
RESULT 5
US-10-421-684-1170
; Sequence 1170, Application US/10421684
; GENERAL INFORMATION:
; APPLICANT: Bunkers, Greg J.
```

```
; APPLICANT: Liang, Jihong
; APPLICANT: Miltanck, Cindy A.
; APPLICANT: Seale, Jeffrey W.
```

```
; APPLICANT: Wu, Yonnie S.
; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
```

```
; FILE REFERENCE: 38-21(15036)B
; CURRENT APPLICATION NUMBER: US/10/421,684
; CURRENT FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: US/09/732,210
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,513
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: US 60/169,340
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 1753
; SEQ ID NO 1170
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Marchantia polymorpha
US-10-421-684-1170
```

```
Query Match          69.2%; Score 18; DB 6; Length 100;
Best Local Similarity 40.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
OY 1 KKKXXXXXLD 10
    ||
DB 29 KRIETSSLD 38
```

```
RESULT 6
US-09-724-676-55061
; Sequence 55061, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 55061
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-55061
```

```
Query Match          69.2%; Score 18; DB 5; Length 109;
Best Local Similarity 40.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
OY 1 KKKXXXXXLD 10
    ||
DB 44 KKSAAESMLD 53
```

```
RESULT 7
US-09-724-676A-55061
; Sequence 55061, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
```

SOFTWARE: PatentIn version 3.2  
SEQ ID NO 55061  
LENGTH: 109  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-724-676A-55061

Query Match 69.2%; Score 18; DB 5; Length 109;  
Best Local Similarity 40.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXXXXXXXXLD 10  
DB 44 KKSAAESMLD 53

RESULT 8  
US-09-513-999C-6288  
Sequence 6288, Application US/09513999C  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J. B.  
APPLICANT: Duclert, A. Y.  
APPLICANT: Giordano, J. Y.  
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
FILE REFERENCE: 59.052.REG  
CURRENT APPLICATION NUMBER: US/09/513,999C  
CURRENT FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/122,487  
PRIOR FILING DATE: 1999-02-26  
NUMBER OF SEQ ID NOS: 36681  
SOFTWARE: Patent.pm  
SEQ ID NO 6288  
LENGTH: 147  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-513-999C-6288

Query Match 69.2%; Score 18; DB 5; Length 147;  
Best Local Similarity 40.0%; Pred. No. 1.8e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXXXXXXXXLD 10  
DB 78 KKKRAATLTD 87

RESULT 9  
US-09-513-999C-6288  
Sequence 6288, Application US/09513999C  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J. B.  
APPLICANT: Duclert, A. Y.  
APPLICANT: Giordano, J. Y.  
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
FILE REFERENCE: 59.052.REG  
CURRENT APPLICATION NUMBER: US/09/513,999C  
CURRENT FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/122,487  
PRIOR FILING DATE: 1999-02-26  
NUMBER OF SEQ ID NOS: 36681  
SOFTWARE: Patent.pm  
SEQ ID NO 6288  
LENGTH: 147  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-513-999C-6288

Query Match 69.2%; Score 18; DB 5; Length 147;  
Best Local Similarity 40.0%; Pred. No. 1.8e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXXXXXXXXLD 10  
DB 78 KKKRAATLTD 87

Db 78 KKKRAATLTD 87

RESULT 10  
US-09-949-016-6460  
Sequence 6460, Application US/09949016  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: C0001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6460  
LENGTH: 152  
TYPE: PRT  
ORGANISM: Human  
US-09-949-016-6460

Query Match 69.2%; Score 18; DB 5; Length 152;  
Best Local Similarity 40.0%; Pred. No. 1.9e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXXXXXXXXLD 10  
DB 44 KKSAAESMLD 53

RESULT 11  
US-60-440-068-114  
Sequence 114, Application US/60440068  
GENERAL INFORMATION:  
APPLICANT: NADLER, STEVEN G.  
APPLICANT: CARMAN, JULIE  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE  
TITLE OF INVENTION: NF-KB PATHWAY  
FILE REFERENCE: 3053-4191  
CURRENT APPLICATION NUMBER: US/60/440,068  
CURRENT FILING DATE: 2003-01-14  
NUMBER OF SEQ ID NOS: 746  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 114  
LENGTH: 152  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-60-440-068-114

Query Match 69.2%; Score 18; DB 7; Length 152;  
Best Local Similarity 40.0%; Pred. No. 1.9e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXXXXXXXXLD 10  
DB 44 KKSAAESMLD 53

RESULT 12  
US-09-724-676-55060  
Sequence 55060, Application US/09724676  
GENERAL INFORMATION:  
APPLICANT: Compugen LTD  
TITLE OF INVENTION: Variants of alternative splicing  
FILE REFERENCE: 129181.4 Compugen  
CURRENT APPLICATION NUMBER: US/09/724,676  
CURRENT FILING DATE: 2000-11-28

NUMBER OF SEQ ID NOS: 97222  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 55060  
LENGTH: 161  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-724-676-55060

Query Match 69.2% Score 18; DB 5; Length 161;  
Best Local Similarity 40.0%; Pred. No. 2e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXXXXXXXXXD 10  
DB 44 KKSNAESMLD 53

## RESULT 13

US-09-724-676A-55060  
Sequence 55060, Application US/09724676A  
GENERAL INFORMATION:  
APPLICANT: Compugen LTD  
TITLE OF INVENTION: Variants of alternative splicing  
FILE REFERENCE: 129181.4 Compugen  
CURRENT APPLICATION NUMBER: US/09/724,676A  
CURRENT FILING DATE: 2000-11-28  
NUMBER OF SEQ ID NOS: 97222  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 55060  
LENGTH: 161  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-724-676A-55060

Query Match 69.2% Score 18; DB 5; Length 161;  
Best Local Similarity 40.0%; Pred. No. 2e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXXXXXXXXXD 10  
DB 44 KKSNAESMLD 53

## RESULT 14

US-10-282-122A-52371  
Sequence 52371, Application US/10282122A  
GENERAL INFORMATION:  
APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trivick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.03A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 52371  
LENGTH: 252  
TYPE: PRT  
ORGANISM: Clostridium botulinum  
US-10-282-122A-52371

Query Match 69.2% Score 18; DB 6; Length 252;  
Best Local Similarity 40.0%; Pred. No. 3.2e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXXXXXXXXXD 10  
DB 9 KRASTDTNLD 18

## RESULT 15

US-10-431-652-7651  
Sequence 7651, Application US/10431652  
GENERAL INFORMATION:  
APPLICANT: Breton, Gary L.  
APPLICANT: Bush, David  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTE  
FILE REFERENCE: PAT03-08  
CURRENT APPLICATION NUMBER: US/10/431,652  
CURRENT FILING DATE: 2003-05-06  
PRIOR APPLICATION NUMBER: US 09/328,352  
PRIOR FILING DATE: 1999-06-04  
PRIOR APPLICATION NUMBER: US 60/088,701  
PRIOR FILING DATE: 1998-06-09  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 7651  
LENGTH: 254  
TYPE: PRT  
ORGANISM: Acinetobacter baumannii  
US-10-431-652-7651

Query Match 69.2% Score 18; DB 6; Length 254;  
Best Local Similarity 40.0%; Pred. No. 3.2e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXXXXXXXXXD 10  
DB 210 KKSXSKTLD 219

Search completed: June 5, 2003, 15:52:26  
Job time : 33.4848 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:29:17 ; Search time 123.182 Seconds

(without alignments)  
52.340 Million cell updates/sec

Title: US-09-150-947F-15

Perfect score: 15

Sequence: 1 XXXXXXXXXD 10

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US097\_COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US099\_COMB.pep.\*  
24: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*  
25: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*  
26: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*  
27: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40.0	2	1	PCT-US00-00358-13	Sequence 13, Appl
2	40.0	2	1	PCT-US00-08879-16	Sequence 16, Appl
3	40.0	2	1	PCT-US00-08879-18	Sequence 18, Appl
4	40.0	2	1	PCT-US00-08879-30	Sequence 30, Appl
5	40.0	2	1	PCT-US00-08879A-16	Sequence 16, Appl
6	40.0	2	1	PCT-US00-08879A-18	Sequence 18, Appl

7	40.0	2	1	PCT-US00-08879A-30	Sequence 30, Appl
8	40.0	2	1	PCT-US02-09188-532	Sequence 532, App
9	40.0	2	1	PCT-US02-09239-152	Sequence 152, App
10	40.0	2	1	PCT-US02-09370-551	Sequence 551, App
11	40.0	2	1	PCT-US02-23763-296	Sequence 296, App
12	40.0	2	1	PCT-US02-30452-85	Sequence 85, Appl
13	40.0	2	1	PCT-US99-26100-6	Sequence 6, Appl1
14	40.0	2	3	US-07-853-042A-3	Sequence 3, Appl1
15	40.0	2	6	US-08-222-626A-122	Sequence 122, App
16	40.0	2	6	US-08-248-816-8	Sequence 8, Appl1
17	40.0	2	8	US-08-430-121-8	Sequence 8, Appl1
18	40.0	2	8	US-08-472-679B-15	Sequence 15, Appl
19	40.0	2	8	US-08-472-679B-23	Sequence 23, Appl
20	40.0	2	8	US-08-472-679B-33	Sequence 33, Appl
21	40.0	2	8	US-08-472-679B-39	Sequence 39, Appl
22	40.0	2	8	US-08-472-679D-15	Sequence 15, Appl
23	40.0	2	8	US-08-472-679D-23	Sequence 23, Appl
24	40.0	2	8	US-08-472-679D-33	Sequence 33, Appl
25	40.0	2	8	US-08-472-679D-39	Sequence 39, Appl
26	40.0	2	8	US-08-472-679H-15	Sequence 15, Appl
27	40.0	2	8	US-08-472-679H-23	Sequence 23, Appl
28	40.0	2	8	US-08-472-679H-33	Sequence 33, Appl
29	40.0	2	8	US-08-472-679H-39	Sequence 39, Appl
30	40.0	2	10	US-08-633-879B-85	Sequence 85, Appl
31	40.0	2	11	US-08-783-79B-85	Sequence 85, Appl
32	40.0	2	12	US-08-843-675A-72	Sequence 72, Appl
33	40.0	2	16	US-09-295-846-24	Sequence 24, Appl
34	40.0	2	16	US-09-295-846-26	Sequence 26, Appl
35	40.0	2	16	US-09-295-846-38	Sequence 38, Appl
36	40.0	2	16	US-09-295-846A-24	Sequence 24, Appl
37	40.0	2	16	US-09-295-846A-26	Sequence 26, Appl
38	40.0	2	16	US-09-295-846A-38	Sequence 38, Appl
39	40.0	2	16	US-09-295-846A-24	Sequence 24, Appl
40	40.0	2	16	US-09-295-846B-26	Sequence 26, Appl
41	40.0	2	16	US-09-295-846B-38	Sequence 38, Appl
42	40.0	2	16	US-09-295-924B-18	Sequence 18, Appl
43	40.0	2	16	US-09-295-924B-20	Sequence 20, Appl
44	40.0	2	16	US-09-295-924B-32	Sequence 32, Appl
45	40.0	2	16	US-09-296-113A-21	Sequence 21, Appl

## ALIGNMENTS

RESULT 1

PCT-US00-00358-13

Sequence 13, Application PC/TUS0000358

GENERAL INFORMATION:

APPLICANT: pecker, Iris

TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY DISTANTLY HOW

FILE REFERENCE: 00/20105

CURRENT APPLICATION NUMBER: PCT/US00/00358

CURRENT FILING DATE: 2000-01-07

PRIOR APPLICATION NUMBER: 60/140,801

PRIOR FILING DATE: 1999-06-25

NUMBER OF SEQ ID NOS: 25

SOFTWARE: PatentIn version 3.0

SEQ ID NO 13

LENGTH: 2

TYPE: PRT

ORGANISM: Homo sapiens

PCT-US00-00358-13

Query Match 40.0%; Score 6; DB 1; Length 2;

Best Local Similarity 100.0%; Pred. No. 4.2e+06;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10

DB 1 D 1

RESULT 2  
PCT-US00-08879-16  
; Sequence 16, Application PC/TUS0008879  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, John  
; APPLICANT: Brandt, Alan  
; APPLICANT: Borovski, Dov  
; TITLE OF INVENTION: Compositions and Methods for Controlling Pests  
; FILE REFERENCE: 4137-120 PCT  
; CURRENT APPLICATION NUMBER: PCT/US00/08879  
; CURRENT FILING DATE: 2000-04-04  
; PRIOR FILING DATE: 1999-04-21  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: TMOF Analogue  
PCT-US00-08879-16

Query Match 40.0%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
|  
Db 1 D 1

RESULT 3  
PCT-US00-08879-18  
; Sequence 18, Application PC/TUS0008879  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, John  
; APPLICANT: Brandt, Alan  
; APPLICANT: Borovski, Dov  
; TITLE OF INVENTION: Compositions and Methods for Controlling Pests  
; FILE REFERENCE: 4137-120 PCT  
; CURRENT APPLICATION NUMBER: PCT/US00/08879  
; CURRENT FILING DATE: 2000-04-04  
; PRIOR APPLICATION NUMBER: 09/295,924  
; PRIOR FILING DATE: 1999-04-21  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: TMOF Analogue  
PCT-US00-08879-18

Query Match 40.0%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
|  
Db 1 D 1

RESULT 4  
PCT-US00-08879-30  
; Sequence 30, Application PC/TUS0008879  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, John  
; APPLICANT: Brandt, Alan  
; APPLICANT: Borovski, Dov  
; TITLE OF INVENTION: Compositions and Methods for Controlling Pests  
; FILE REFERENCE: 4137-120 PCT  
; OTHER INFORMATION: Description of Artificial Sequence: TMOF Analogue  
PCT-US00-08879-18

; CURRENT APPLICATION NUMBER: PCT/US00/08879  
; CURRENT FILING DATE: 2000-04-04  
; PRIOR APPLICATION NUMBER: 09/295,924  
; PRIOR FILING DATE: 1999-04-21  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 30  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: TMOF Analogue  
PCT-US00-08879-30

Query Match 40.0%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
|  
Db 2 D 2

RESULT 5  
PCT-US00-08879A-16  
; Sequence 16, Application PC/TUS0008879A  
; GENERAL INFORMATION:  
; APPLICANT: Insect Biotechnology, Inc.  
; TITLE OF INVENTION: Compositions and Methods for Controlling Pests  
; FILE REFERENCE: 4137-120 PCT  
; CURRENT APPLICATION NUMBER: PCT/US00/08879A  
; CURRENT FILING DATE: 2000-04-04  
; PRIOR APPLICATION NUMBER: 09/295,924  
; PRIOR FILING DATE: 1999-04-21  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: TMOF Analogue  
PCT-US00-08879A-16

Query Match 40.0%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
|  
Db 1 D 1

RESULT 6  
PCT-US00-08879A-18  
; Sequence 18, Application PC/TUS0008879A  
; GENERAL INFORMATION:  
; APPLICANT: Insect Biotechnology, Inc.  
; TITLE OF INVENTION: Compositions and Methods for Controlling Pests  
; FILE REFERENCE: 4137-120 PCT  
; CURRENT APPLICATION NUMBER: PCT/US00/08879A  
; CURRENT FILING DATE: 2000-04-04  
; PRIOR APPLICATION NUMBER: 09/295,924  
; PRIOR FILING DATE: 1999-04-21  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: TMOF Analogue  
PCT-US00-08879A-18

Query Match 40.0%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
DB 1 D 1

## RESULT 7

PCT-US00-08879A-30  
; Sequence 30, Application PC/TUS0008879A  
; GENERAL INFORMATION:  
; APPLICANT: Insect Biotechnology, Inc.  
; TITLE OF INVENTION: Compositions and Methods for Controlling Pests  
; FILE REFERENCE: 4137-120 PCT  
; CURRENT APPLICATION NUMBER: PCT/US00/08879A  
; PRIOR FILING DATE: 2000-04-04  
; PRIOR APPLICATION NUMBER: 09/295,924  
; PRIOR FILING DATE: 1999-04-21  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 30  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: TMOF Analogue  
PCT-US00-08879A-30

Query Match 40.0%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
DB 2 D 2

## RESULT 8

PCT-US02-09188-532  
; Sequence 532, Application PC/TUS0209188  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Human Secreted Proteins  
; FILE REFERENCE: PS952PCT  
; CURRENT APPLICATION NUMBER: PCT/US02/09188  
; CURRENT FILING DATE: 2002-03-26  
; PRIOR APPLICATION NUMBER: US 60/278,650  
; PRIOR FILING DATE: 2001-03-27  
; PRIOR APPLICATION NUMBER: US 09/950,082  
; PRIOR FILING DATE: 2001-09-12  
; PRIOR APPLICATION NUMBER: US 09/950,083  
; PRIOR FILING DATE: 2001-09-12  
; NUMBER OF SEQ ID NOS: 1732  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 532  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US02-09188-532

Query Match 40.0%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
DB 1 D 1

## RESULT 9

PCT-US02-09239-152  
; Sequence 152, Application PC/TUS0209239  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Human Secreted Proteins  
; FILE REFERENCE: PS953PCT  
; CURRENT APPLICATION NUMBER: PCT/US02/09239  
; CURRENT FILING DATE: 2002-03-26  
; PRIOR APPLICATION NUMBER: US 60/278,650  
; PRIOR FILING DATE: 2001-03-27  
; PRIOR APPLICATION NUMBER: US 09/950,082  
; PRIOR FILING DATE: 2001-09-12  
; PRIOR APPLICATION NUMBER: US 09/950,083  
; PRIOR FILING DATE: 2001-09-12  
; NUMBER OF SEQ ID NOS: 380  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 152  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US02-09239-152

Query Match 40.0%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
DB 1 D 1

## RESULT 10

PCT-US02-09370-551  
; Sequence 551, Application PC/TUS0209370  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Human Secreted Proteins  
; FILE REFERENCE: PS954PCT  
; CURRENT APPLICATION NUMBER: PCT/US02/09370  
; CURRENT FILING DATE: 2002-03-26  
; PRIOR APPLICATION NUMBER: US 60/278,650  
; PRIOR FILING DATE: 2001-03-27  
; PRIOR APPLICATION NUMBER: US 09/950,082  
; PRIOR FILING DATE: 2001-09-12  
; PRIOR APPLICATION NUMBER: US 09/950,083  
; PRIOR FILING DATE: 2001-09-12  
; NUMBER OF SEQ ID NOS: 1834  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 551  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US02-09370-551

Query Match 40.0%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
DB 1 D 1

## RESULT 11

PCT-US02-23763-296  
; Sequence 296, Application PC/TUS0223763  
; GENERAL INFORMATION:  
; APPLICANT: Sundaramoorthy, M.  
; TITLE OF INVENTION: Crystallized structure of Type IV Collagen NCI Domain Hexamer  
; FILE REFERENCE: MBHB 01-1017  
; CURRENT APPLICATION NUMBER: PCT/US02/23763  
; CURRENT FILING DATE: 2002-07-26

PRIOR APPLICATION NUMBER: US 60/308,523  
PRIOR FILING DATE: 2001-07-27  
PRIOR APPLICATION NUMBER: US 60/351,289  
PRIOR FILING DATE: 2001-10-29  
PRIOR APPLICATION NUMBER: US 60/366,854  
PRIOR FILING DATE: 2002-03-22  
PRIOR APPLICATION NUMBER: US 60/385,362  
PRIOR FILING DATE: 2002-06-03  
NUMBER OF SEQ ID NOS: 307  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 296  
LENGTH: 2  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US02-23763-296

Query Match 40.0%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
Db 2 D 2

RESULT 12  
PCT-US02-30452-85  
Sequence 85, Application PC/TUS0230452  
GENERAL INFORMATION:  
APPLICANT: Wyeth  
TITLE OF INVENTION: Antibody Inhibitors of GDF-8 and Uses Therefor  
FILE REFERENCE: 08702.00012-00304  
CURRENT APPLICATION NUMBER: PCT/US02/30452  
CURRENT FILING DATE: 2002-09-26  
PRIOR APPLICATION NUMBER: 60/334,528  
PRIOR FILING DATE: 2001-09-26  
NUMBER OF SEQ ID NOS: 131  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 85  
LENGTH: 2  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US02-30452-85

Query Match 40.0%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
Db 2 D 2

RESULT 13  
PCT-US99-26100-6  
Sequence 6, Application PC/TUS9926100  
GENERAL INFORMATION:  
APPLICANT: Ream, Walt et al.,  
TITLE OF INVENTION: Plants Having Enhanced Gall Resistance and  
TITLE OF INVENTION: Methods and Compositions for Producing the Same  
FILE REFERENCE: 53629  
CURRENT APPLICATION NUMBER: PCT/US99/26100  
CURRENT FILING DATE: 1999-11-04  
EARLIER APPLICATION NUMBER: 60/107,185  
EARLIER FILING DATE: 1998-11-05  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 6  
LENGTH: 2  
TYPE: PRT  
ORGANISM: Agrobacterium tumefaciens  
PCT-US99-26100-6

Query Match 40.0%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
Db 2 D 2

RESULT 14  
US-07-853-042A-3  
Sequence 3, Application US/07853042A  
GENERAL INFORMATION:  
APPLICANT: Anders BIL  
TITLE OF INVENTION: METHOD AND MEANS FOR INDUCING, RESP.,  
TITLE OF INVENTION: PREVENTING CONSTRICTION OF THE PUPIL  
TITLE OF INVENTION: IN THE EYE  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FRED PHILPITT  
STREET: 99 Canal Center Plaza, Ste. 300  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22314

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: DOS Text File  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/853,042A  
FILING DATE: 19920318  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: FRED PHILPITT  
REGISTRATION NUMBER: 16825  
REFERENCE/DOCKET NUMBER: PHE-808  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703 684 1111  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Polypeptide

US-07-853-042A-3

Query Match 40.0%; Score 6; DB 3; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
Db 1 D 1

RESULT 15  
US-08-222-626A-122  
Sequence 122, Application US/08222626A  
GENERAL INFORMATION:  
APPLICANT: Aral, Naoko  
APPLICANT: Masuda, Esteban S.  
APPLICANT: Tokumitsu, Hiroshi  
TITLE OF INVENTION: PURIFIED COMPONENTS OF MAMMALIAN  
TITLE OF INVENTION: TRANSCRIPTION REGULATION COMPLEXES, AND ANALOGS  
NUMBER OF SEQUENCES: 183  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DNAX Research Institute  
STREET: 901 California Avenue  
CITY: Palo Alto  
STATE: California

COUNTRY: USA  
 ZIP: 94304-1104  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/222,626A  
 FILING DATE: 04-APR-1994  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/148,061  
 FILING DATE: 05-NOV-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/113,971  
 FILING DATE: 30-AUG-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/088,998  
 FILING DATE: 30-JUL-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/088,483  
 FILING DATE: 06-JUL-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Ching, Edwin P.  
 REGISTRATION NUMBER: 34,090  
 REFERENCE/DOCKET NUMBER: DX0392K3  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-852-9196  
 TELEFAX: 415-496-1200  
 INFORMATION FOR SEQ ID NO: 122:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-222-626A-122

Query Match 40.0%; Score 6; DB 6; Length 2;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 10 D 10  
 Db 1 D 1

Search completed: June 5, 2003, 15:47:37  
 Job time : 124.182 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:31:47 ; Search time 28.4848 seconds  
(without alignments)  
74.544 Million cell updates/sec

Title: US-09-150-947f-15  
Perfect score: 15  
Sequence: 1 xxxxxxxxxxXD 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1012620 seqs, 212337834 residues

Total number of hits satisfying chosen parameters: 1012620

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*  
1: /cgn2\_6/ptodata/1/paa/pct\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	40.0	2	4	US-08-472-679H-15
2	6	40.0	2	4	US-08-472-679H-23
3	6	40.0	2	4	US-08-472-679H-33
4	6	40.0	2	4	US-08-472-679H-39
5	6	40.0	2	4	US-08-472-679H-15
6	6	40.0	2	4	US-08-472-679H-23
7	6	40.0	2	4	US-08-472-679H-33
8	6	40.0	2	4	US-08-472-679H-39
9	6	40.0	2	6	US-10-109-048-461
10	6	40.0	2	6	US-10-109-048-461
11	6	40.0	2	6	US-10-405-027-5115
12	6	40.0	3	1	PCT-US02-14358-15
13	6	40.0	3	1	PCT-US02-14358-16
14	6	40.0	3	1	PCT-US02-14358-17
15	6	40.0	3	1	PCT-US02-14358-18
16	6	40.0	3	1	PCT-US02-14358-19
17	6	40.0	3	1	PCT-US02-14358-20
18	6	40.0	3	1	PCT-US02-14358-20
19	6	40.0	3	1	PCT-US02-14358-20
20	6	40.0	3	1	PCT-US02-14358-20
21	6	40.0	3	1	PCT-US02-14358-20
22	6	40.0	3	1	PCT-US02-14358-20
23	6	40.0	3	1	PCT-US02-14358-20
24	6	40.0	3	1	PCT-US02-14358-20
25	6	40.0	3	1	PCT-US02-14358-20
26	6	40.0	3	1	PCT-US02-14358-20

27	6	40.0	3	1	PCT-US03-14096-24	Sequence 24, Appl
28	6	40.0	3	5	US-09-302-195C-1	Sequence 1, Appl
29	6	40.0	3	5	US-09-302-195C-1	Sequence 11, Appl
30	6	40.0	3	5	US-09-716-394-11	Sequence 37, Appl
31	6	40.0	3	5	US-09-853-080A-37	Sequence 11, Appl
32	6	40.0	3	5	US-09-772-819-11	Sequence 11, Appl
33	6	40.0	3	5	US-09-573-830-11	Sequence 11, Appl
34	6	40.0	3	5	US-09-992-124B-74	Sequence 74, Appl
35	6	40.0	3	5	US-09-564-045-11	Sequence 11, Appl
36	6	40.0	3	5	US-09-867-847A-30	Sequence 30, Appl
37	6	40.0	3	5	US-09-980-347-6	Sequence 6, Appl
38	6	40.0	3	5	US-09-723-437B-11	Sequence 11, Appl
39	6	40.0	3	5	US-09-939-481-81	Sequence 81, Appl
40	6	40.0	3	5	US-09-898-165B-33	Sequence 33, Appl
41	6	40.0	3	5	US-09-623-548A-1053	Sequence 1053, Ap
42	6	40.0	3	5	US-09-623-548A-1054	Sequence 1054, Ap
43	6	40.0	3	5	US-09-623-548A-1160	Sequence 1160, Ap
44	6	40.0	3	6	US-09-623-548A-1170	Sequence 1170, Ap
45	6	40.0	3	6	US-10-141-531-15	Sequence 15, Appl
					US-10-141-531-16	Sequence 16, Appl

## ALIGNMENTS

RESULT 1  
US-08-472-679H-15  
Sequence 15, Application US/08472679H  
GENERAL INFORMATION:

APPLICANT: Cochran, Mark D.  
Junker, David E.

TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 267  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Pamela G. Salkeld  
STREET: 2000 Galloping Hill Road  
CITY: Kenilworth  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07033

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/472,679H  
FILING DATE: 07-Jun-1995

CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:

NAME: Salkeld, Pamela G.  
REGISTRATION NUMBER: 38,607  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 298-2135  
TELEFAX: (908) 298-5388

INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:

LENGTH: 2 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 15:

US-08-472-679H-15

Query Match 40.0%; Score 6; DB 4; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 D 10  
Db 1 D 1

RESULT 2

US-08-472-679H-23  
; Sequence 23, Application US/08472679H  
; GENERAL INFORMATION:  
; APPLICANT: Cochran, Mark D.  
; Junker, David E.  
; TITLE OF INVENTION: Recombinant Swinepox Virus  
; NUMBER OF SEQUENCES: 267  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pamela G. Salkeld  
; STREET: 2000 Galloping Hill Road  
; CITY: Kenilworth  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07033  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472,679H  
; FILING DATE: 07-Jun-1995  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Salkeld, Pamela G.  
; REGISTRATION NUMBER: 38,607  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908) 298-2135  
; TELEFAX: (908) 298-5388  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:  
US-08-472-679H-23  
Query Match 40.0%; Score 6; DB 4; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 10 D 10  
Db 1 D 1  
RESULT 3  
US-08-472-679H-33  
; Sequence 33, Application US/08472679H  
; GENERAL INFORMATION:  
; APPLICANT: Cochran, Mark D.  
; Junker, David E.  
; TITLE OF INVENTION: Recombinant Swinepox Virus  
; NUMBER OF SEQUENCES: 267  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pamela G. Salkeld  
; STREET: 2000 Galloping Hill Road  
; CITY: Kenilworth  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07033  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472,679H  
; FILING DATE: 07-Jun-1995  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Salkeld, Pamela G.  
; REGISTRATION NUMBER: 38,607  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908) 298-2135  
; TELEFAX: (908) 298-5388  
; INFORMATION FOR SEQ ID NO: 33:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 33:  
US-08-472-679H-33  
Query Match 40.0%; Score 6; DB 4; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 10 D 10  
Db 1 D 1  
RESULT 4  
US-08-472-679H-39  
; Sequence 39, Application US/08472679H  
; GENERAL INFORMATION:  
; APPLICANT: Cochran, Mark D.  
; Junker, David E.  
; TITLE OF INVENTION: Recombinant Swinepox Virus  
; NUMBER OF SEQUENCES: 267  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pamela G. Salkeld  
; STREET: 2000 Galloping Hill Road  
; CITY: Kenilworth  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07033  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472,679H  
; FILING DATE: 07-Jun-1995  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Salkeld, Pamela G.  
; REGISTRATION NUMBER: 38,607  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908) 298-2135  
; TELEFAX: (908) 298-5388  
; INFORMATION FOR SEQ ID NO: 39:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 39:  
US-08-472-679H-39  
Query Match 40.0%; Score 6; DB 4; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 10 D 10  
Db 1 D 1  
RESULT 5  
US-08-472-679H-15  
; Sequence 15, Application US/08472679H

US-08-472-679H-39  
; Sequence 39, Application US/08472679H  
; GENERAL INFORMATION:  
; APPLICANT: Cochran, Mark D.  
; Junker, David E.  
; TITLE OF INVENTION: Recombinant Swinepox Virus  
; NUMBER OF SEQUENCES: 267  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pamela G. Salkeld  
; STREET: 2000 Galloping Hill Road  
; CITY: Kenilworth  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07033  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472,679H  
; FILING DATE: 07-Jun-1995  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Salkeld, Pamela G.  
; REGISTRATION NUMBER: 38,607  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908) 298-2135  
; TELEFAX: (908) 298-5388  
; INFORMATION FOR SEQ ID NO: 39:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 39:  
US-08-472-679H-39  
Query Match 40.0%; Score 6; DB 4; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 10 D 10  
Db 1 D 1  
RESULT 5  
US-08-472-679H-15  
; Sequence 15, Application US/08472679H



GENERAL INFORMATION:  
APPLICANT: Cochran, Mark D.  
Junker, David E.  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 267  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pamela G. Salkeld  
STREET: 2000 Galloping Hill Road  
CITY: Kenilworth  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07033  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,679H  
FILING DATE: 07-Jun-1995  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Salkeld, Pamela G.  
REGISTRATION NUMBER: 38,607  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 298-2135  
TELEFAX: (908) 298-5388  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
MOLECULE TYPE: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 15:  
US-08-472-679H-15

Query Match 40.0%; Score 6; DB 4; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
DB 1 D 1

RESULT 6  
US-08-472-679H-23  
Sequence 23, Application US/08472679H  
GENERAL INFORMATION:  
APPLICANT: Cochran, Mark D.  
Junker, David E.  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 267  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pamela G. Salkeld  
STREET: 2000 Galloping Hill Road  
CITY: Kenilworth  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07033  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,679H  
FILING DATE: 07-Jun-1995  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Salkeld, Pamela G.  
REGISTRATION NUMBER: 38,607  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 298-2135  
TELEFAX: (908) 298-5388  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
MOLECULE TYPE: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 33:  
US-08-472-679H-33

TELEPHONE: (908) 298-2135  
TELEFAX: (908) 298-5388  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
MOLECULE TYPE: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 23:  
US-08-472-679H-23

Query Match 40.0%; Score 6; DB 4; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
DB 1 D 1

RESULT 7  
US-08-472-679H-33  
Sequence 33, Application US/08472679H  
GENERAL INFORMATION:  
APPLICANT: Cochran, Mark D.  
Junker, David E.  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 267  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pamela G. Salkeld  
STREET: 2000 Galloping Hill Road  
CITY: Kenilworth  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07033  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,679H  
FILING DATE: 07-Jun-1995  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Salkeld, Pamela G.  
REGISTRATION NUMBER: 38,607  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 298-2135  
TELEFAX: (908) 298-5388  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
MOLECULE TYPE: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 33:  
US-08-472-679H-33

Query Match 40.0%; Score 6; DB 4; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 D 10  
DB 1 D 1

RESULT 8  
US-08-472-679H-39  
Sequence 39, Application US/08472679H  
GENERAL INFORMATION:  
APPLICANT: Cochran, Mark D.

Junker, David E.  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 267  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pamela G. Salkeld  
CITY: Kenilworth  
STREET: 2000 Galloping Hill Road  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07033  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,679H  
FILING DATE: 07-Jun-1995  
CLASSIFICATION: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Salkeld, Pamela G.  
REGISTRATION NUMBER: 38,607  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 298-2135  
TELEFAX: (908) 298-5388  
INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 39:  
US-08-472-679H-39

Query Match 40.0%; Score 6; DB 4; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
DB 1 D 1

RESULT 9  
US-10-187-339-6  
; Sequence 6, Application US/10187339  
; GENERAL INFORMATION:  
; APPLICANT: Ream, Walt et al.  
; TITLE OF INVENTION: Plants Having Enhanced Gall Resistance and  
; FILE OF INVENTION: Methods and Compositions for Producing the Same  
; FILE REFERENCE: 53629  
; CURRENT APPLICATION NUMBER: US/10/187,339  
; CURRENT FILING DATE: 2002-06-28  
; PRIOR APPLICATION NUMBER: US/09/434,837  
; PRIOR FILING DATE: 1999-11-04  
; PRIOR APPLICATION NUMBER: 60/107,185  
; PRIOR FILING DATE: 1998-11-05  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Agrobacterium tumefaciens  
US-10-187-339-6

Query Match 40.0%; Score 6; DB 6; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
DB 2 D 2

RESULT 10  
US-10-109-048-461  
; Sequence 461, Application US/10109048  
; GENERAL INFORMATION:  
; APPLICANT: COMMORI, PADMA  
; APPLICANT: KEELING, PETER L.  
; APPLICANT: RAMIREZ, NONA  
; APPLICANT: MCKEAN, ANGELA  
; APPLICANT: GAO, ZHONG  
; APPLICANT: GUAN, HANPING  
; TITLE OF INVENTION: GLUCAN CHAIN LENGTH DOMAINS  
; FILE REFERENCE: 2461-76  
; CURRENT APPLICATION NUMBER: US/10/109,048  
; CURRENT FILING DATE: 2003-03-04  
; PRIOR APPLICATION NUMBER: 60/279,720  
; PRIOR FILING DATE: 2001-03-30  
; NUMBER OF SEQ ID NOS: 1154  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 461  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: Accession No. 3832512  
US-10-109-048-461

Query Match 40.0%; Score 6; DB 6; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
DB 2 D 2

RESULT 11  
US-10-405-027-5115  
; Sequence 5115, Application US/10405027  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Human Secreted Proteins  
; FILE REFERENCE: PS806P1  
; CURRENT APPLICATION NUMBER: US/10/405,027  
; CURRENT FILING DATE: 2003-04-07  
; PRIOR APPLICATION NUMBER: 60/369,608  
; PRIOR FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: 60/376,175  
; PRIOR FILING DATE: 2002-04-30  
; NUMBER OF SEQ ID NOS: 5810  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 5115  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-405-027-5115

Query Match 40.0%; Score 6; DB 6; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
DB 1 D 1

RESULT 12  
PCT-US02-14358-15  
; Sequence 15, Application PC/TUS0214358  
; GENERAL INFORMATION:  
; APPLICANT: XENCOR  
; APPLICANT: SINGENTA PARTICIPATIONS AG  
; APPLICANT: Briggs, Steven P.

APPLICANT: Dalmia, Bipin K.  
APPLICANT: del Val, Greg  
APPLICANT: Desjarlais, John R.  
APPLICANT: Helfetz, Peter  
APPLICANT: Lugnbuhl, Peter  
APPLICANT: Muchhal, Umesh  
TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity  
FILE REFERENCE: FP-71457-2-PC/RT/RMS/RMK  
CURRENT APPLICATION NUMBER: PCT/US02/14358  
PRIOR FILING DATE: 2002-11-25  
PRIOR APPLICATION NUMBER: US 60/370,609  
PRIOR FILING DATE: 2002-04-05  
PRIOR APPLICATION NUMBER: US 60/289,029  
PRIOR FILING DATE: 2001-05-04  
NUMBER OF SEQ ID NOS: 208  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 15  
LENGTH: 3  
TYPE: PRT  
ORGANISM: Escherichia coli  
PCT-US02-14358-15

Query Match 40.0%; Score 6; DB 1; Length 3;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
|  
Db 1 D 1

RESULT 13  
PCT-US02-14358-16  
Sequence 16, Application PC/TUS0214358  
GENERAL INFORMATION:  
APPLICANT: XENCOR  
APPLICANT: SYNGENTA PARTICIPATIONS AG  
APPLICANT: Briggs, Steven P.  
APPLICANT: Dalmia, Bipin K.  
APPLICANT: del Val, Greg  
APPLICANT: Desjarlais, John R.  
APPLICANT: Helfetz, Peter  
APPLICANT: Lugnbuhl, Peter  
APPLICANT: Muchhal, Umesh  
TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity  
FILE REFERENCE: FP-71457-2-PC/RT/RMS/RMK  
CURRENT APPLICATION NUMBER: PCT/US02/14358  
PRIOR FILING DATE: 2002-11-25  
PRIOR APPLICATION NUMBER: US 60/370,609  
PRIOR FILING DATE: 2002-04-05  
PRIOR APPLICATION NUMBER: US 60/289,029  
PRIOR FILING DATE: 2001-05-04  
NUMBER OF SEQ ID NOS: 208  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 16  
LENGTH: 3  
TYPE: PRT  
ORGANISM: Bacillus subtilis  
PCT-US02-14358-16

Query Match 40.0%; Score 6; DB 1; Length 3;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
|  
Db 1 D 1

RESULT 14  
PCT-US02-14358-17  
Sequence 17, Application PC/TUS0214358  
GENERAL INFORMATION:

APPLICANT: XENCOR  
APPLICANT: SYNGENTA PARTICIPATIONS AG  
APPLICANT: Briggs, Steven P.  
APPLICANT: Dalmia, Bipin K.  
APPLICANT: del Val, Greg  
APPLICANT: Desjarlais, John R.  
APPLICANT: Helfetz, Peter  
APPLICANT: Lugnbuhl, Peter  
APPLICANT: Muchhal, Umesh  
TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity  
FILE REFERENCE: FP-71457-2-PC/RT/RMS/RMK  
CURRENT APPLICATION NUMBER: PCT/US02/14358  
PRIOR FILING DATE: 2002-11-25  
PRIOR APPLICATION NUMBER: US 60/370,609  
PRIOR FILING DATE: 2002-04-05  
PRIOR APPLICATION NUMBER: US 60/289,029  
PRIOR FILING DATE: 2001-05-04  
NUMBER OF SEQ ID NOS: 208  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 17  
LENGTH: 3  
TYPE: PRT  
ORGANISM: Mycobacterium leprae  
PCT-US02-14358-17

Query Match 40.0%; Score 6; DB 1; Length 3;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
|  
Db 1 D 1

RESULT 15  
PCT-US02-14358-18  
Sequence 18, Application PC/TUS0214358  
GENERAL INFORMATION:  
APPLICANT: XENCOR  
APPLICANT: SYNGENTA PARTICIPATIONS AG  
APPLICANT: Briggs, Steven P.  
APPLICANT: Dalmia, Bipin K.  
APPLICANT: del Val, Greg  
APPLICANT: Desjarlais, John R.  
APPLICANT: Helfetz, Peter  
APPLICANT: Lugnbuhl, Peter  
APPLICANT: Muchhal, Umesh  
TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity  
FILE REFERENCE: FP-71457-2-PC/RT/RMS/RMK  
CURRENT APPLICATION NUMBER: PCT/US02/14358  
PRIOR FILING DATE: 2002-11-25  
PRIOR APPLICATION NUMBER: US 60/370,609  
PRIOR FILING DATE: 2002-04-05  
PRIOR APPLICATION NUMBER: US 60/289,029  
PRIOR FILING DATE: 2001-05-04  
NUMBER OF SEQ ID NOS: 208  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 18  
LENGTH: 3  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
PCT-US02-14358-18

Query Match 40.0%; Score 6; DB 1; Length 3;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 D 10  
|  
Db 1 D 1

Search completed: June 5, 2003, 15:52:27

Fri Jun 6 09:14:41 2003

us-09-150-947f-15.rapn

Page 6

Job time : 29.4848 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:29:17 ; Search time 147.818 Seconds

(Without alignments)  
52,340 Million cell updates/sec

Title: US-09-150-947F-16

Perfect score: 33

Sequence: 1 XXXXXXTXQEXD 12

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 64473110 residues  
Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US097\_COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US099\_COMB.pep.\*  
24: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*  
25: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*  
26: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*  
27: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	24	72.7	204	US-09-869-136-9	Sequence 9, Appl1
2	24	72.7	209	US-09-791-537-103693	Sequence 103693,
3	24	72.7	209	US-09-791-537-103695	Sequence 103695,
4	24	72.7	209	US-09-791-537-103701	Sequence 103701,
5	24	72.7	209	US-09-791-537-103704	Sequence 103704,
6	24	72.7	209	US-09-791-537-103705	Sequence 103705,

7	24	72.7	209	US-09-791-537-103706	Sequence 103706,
8	24	72.7	209	US-09-791-537-103708	Sequence 103708,
9	24	72.7	209	US-09-791-537-103712	Sequence 103712,
10	24	72.7	209	US-09-791-537-103715	Sequence 103715,
11	24	72.7	209	US-09-791-537-103717	Sequence 103717,
12	24	72.7	209	US-09-791-537-103719	Sequence 103719,
13	24	72.7	209	US-09-791-537-103720	Sequence 103720,
14	24	72.7	209	US-09-791-537-103729	Sequence 103729,
15	24	72.7	209	US-09-791-537-103733	Sequence 103733,
16	24	72.7	209	US-09-791-537-103736	Sequence 103736,
17	24	72.7	209	US-09-791-537-103738	Sequence 103738,
18	24	72.7	209	US-09-791-537-103740	Sequence 103740,
19	24	72.7	209	US-09-791-537-120755	Sequence 120755,
20	24	72.7	209	US-09-791-537-137418	Sequence 137418,
21	24	72.7	209	US-09-791-537-85735	Sequence 85735, A
22	24	72.7	233	US-09-869-136-2	Sequence 2, Appl1
23	24	72.7	256	US-09-791-537-86827	Sequence 86827, A
24	24	72.7	256	US-09-791-537-88724	Sequence 88724, A
25	24	72.7	256	US-09-791-537-97492	Sequence 97492, A
26	24	72.7	278	US-09-675-784A-11175	Sequence 11175, A
27	24	72.7	278	US-09-675-784A-11175	Sequence 11175, A
28	24	72.7	603	US-09-791-537-25751	Sequence 25751, A
29	24	72.7	603	US-09-791-537-41676	Sequence 41676, A
30	24	72.7	604	US-09-791-537-22679	Sequence 22679, A
31	24	72.7	604	US-09-791-537-126252	Sequence 126252,
32	24	72.7	955	US-09-791-537-100507	Sequence 100507,
33	23	69.7	17	US-08-491-746-6	Sequence 6, Appl1
34	23	69.7	17	US-09-555-115A-30	Sequence 30, Appl
35	23	69.7	17	US-09-555-115A-31	Sequence 31, Appl
36	23	69.7	17	US-09-555-115A-32	Sequence 32, Appl
37	23	69.7	17	US-09-555-115A-32	Sequence 32, Appl
38	23	69.7	17	US-09-555-115A-32	Sequence 32, Appl
39	23	69.7	24	PCT-US00-16680A-20	Sequence 20, Appl
40	23	69.7	24	US-08-838-413-20	Sequence 20, Appl
41	23	69.7	24	US-09-335-581-20	Sequence 20, Appl
42	23	69.7	24	US-09-335-581A-20	Sequence 20, Appl
43	23	69.7	24	US-09-336-627B-20	Sequence 20, Appl
44	23	69.7	24	US-09-335-581-20	Sequence 20, Appl
45	23	69.7	29	PCT-US01-00663-32245	Sequence 32245, A
			22	US-09-864-761-43487	Sequence 43487, A

## ALIGNMENTS

RESULT 1  
US-09-869-136-9  
Sequence 9, Application US/09869136  
GENERAL INFORMATION:  
APPLICANT: FRASER, JOHN DAVID  
APPLICANT: PROF, THOMAS  
TITLE OF INVENTION: SUPERANTIGENS  
FILE REFERENCE: 3911-8  
CURRENT APPLICATION NUMBER: US/09/869,136  
CURRENT FILING DATE: 2001-07-20  
PRIOR APPLICATION NUMBER: PCT/NZ99/00228  
PRIOR FILING DATE: 1999-12-24  
PRIOR APPLICATION NUMBER: NZ 333589  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 9  
LENGTH: 204  
TYPE: PRT  
ORGANISM: Streptococcus pyogenes  
US-09-869-136-9

Query Match 72.7%; Score 24; DB 22; Length 204;  
Best Local Similarity 50.0%; Pred. No. 5,9e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 3 XXXXXTQEXD 12  
DB 123 KTTVTAQEXD 132

```
RESULT 2
US-09-791-537-103693
; Sequence 103693, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonmix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 103693
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-09-791-537-103693

Query Match          72.7%; Score 24; DB 21; Length 209;
Best Local Similarity 50.0%; Pred. No. 6.1e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      3 KXXXTXQEXD 12
      | | | | |
Db      125 KTVTAQAEID 134

RESULT 3
US-09-791-537-103695
; Sequence 103695, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonmix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 103695
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-09-791-537-103695

Query Match          72.7%; Score 24; DB 21; Length 209;
Best Local Similarity 50.0%; Pred. No. 6.1e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      3 KXXXTXQEXD 12
      | | | | |
Db      125 KTVTAQAEID 134

RESULT 4
US-09-791-537-103701
; Sequence 103701, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonmix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 103701
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-09-791-537-103701

Query Match          72.7%; Score 24; DB 21; Length 209;
Best Local Similarity 50.0%; Pred. No. 6.1e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      3 KXXXTXQEXD 12
      | | | | |
Db      125 KTVTAQAEID 134

RESULT 5
US-09-791-537-103704
; Sequence 103704, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonmix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 103704
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-09-791-537-103704

Query Match          72.7%; Score 24; DB 21; Length 209;
Best Local Similarity 50.0%; Pred. No. 6.1e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      3 KXXXTXQEXD 12
      | | | | |
Db      125 KTVTAQAEID 134

RESULT 6
US-09-791-537-103705
; Sequence 103705, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonmix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 103705
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-09-791-537-103705

Query Match          72.7%; Score 24; DB 21; Length 209;
Best Local Similarity 50.0%; Pred. No. 6.1e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      3 KXXXTXQEXD 12
      | | | | |
Db      125 KTVTAQAEID 134
```

```

RESULT 7
US-09-791-537-103706
; Sequence 103706, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 103706
; LENGTH: 209
; TYPE: PRF
; ORGANISM: Streptococcus pyogenes
US-09-791-537-103706

```

```

Query Match
Best Local Similarity 72.7%; Score 24; DB 21; Length 209;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

OY 3 KXXXTXQXND 12
Db 125 KTTVTAQXND 134

```

```

RESULT 8
US-09-791-537-103708
; Sequence 103708, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 103708
; LENGTH: 209
; TYPE: PRF
; ORGANISM: Streptococcus pyogenes
US-09-791-537-103708

```

```

Query Match
Best Local Similarity 72.7%; Score 24; DB 21; Length 209;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

OY 3 KXXXTXQXND 12
Db 125 KTTVTAQXND 134

```

```

RESULT 9
US-09-791-537-103712
; Sequence 103712, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0

```

```

; SEQ ID NO 103712
; LENGTH: 209
; TYPE: PRF
; ORGANISM: Streptococcus pyogenes
US-09-791-537-103712

```

```

Query Match
Best Local Similarity 72.7%; Score 24; DB 21; Length 209;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

OY 3 KXXXTXQXND 12
Db 125 KTTVTAQXND 134

```

```

RESULT 10
US-09-791-537-103715
; Sequence 103715, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 103715
; LENGTH: 209
; TYPE: PRF
; ORGANISM: Streptococcus pyogenes
US-09-791-537-103715

```

```

Query Match
Best Local Similarity 72.7%; Score 24; DB 21; Length 209;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

OY 3 KXXXTXQXND 12
Db 125 KTTVTAQXND 134

```

```

RESULT 11
US-09-791-537-103717
; Sequence 103717, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 103717
; LENGTH: 209
; TYPE: PRF
; ORGANISM: Streptococcus pyogenes
US-09-791-537-103717

```

```

Query Match
Best Local Similarity 72.7%; Score 24; DB 21; Length 209;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

OY 3 KXXXTXQXND 12
Db 125 KTTVTAQXND 134

```

```
RESULT 12
US-09-791-537-103719
; Sequence 103719, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 103719
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-09-791-537-103719

Query Match
Best Local Similarity 72.7%; Score 24; DB 21; Length 209;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12
DB 125 KTVYTAQED 134

RESULT 13
US-09-791-537-103720
; Sequence 103720, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 103720
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-09-791-537-103720

Query Match
Best Local Similarity 72.7%; Score 24; DB 21; Length 209;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12
DB 125 KTVYTAQED 134

RESULT 14
US-09-791-537-103729
; Sequence 103729, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 103729
```

```
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-09-791-537-103729

Query Match
Best Local Similarity 72.7%; Score 24; DB 21; Length 209;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12
DB 125 KTVYTAQED 134

RESULT 15
US-09-791-537-103733
; Sequence 103733, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 103733
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-09-791-537-103733

Query Match
Best Local Similarity 72.7%; Score 24; DB 21; Length 209;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12
DB 125 KTVYTAQED 134
```

Search completed: June 5, 2003, 15:47:37  
Job time : 147.818 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:31:47 ; Search time 34.1818 Seconds  
(without alignments)  
74.544 Million cell updates/sec

Title: US-09-150-947F-16  
Perfect score: 33  
Sequence: 1 XXXXXXTXQEXD 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1012620 seqs, 212337834 residues  
Total number of hits satisfying chosen parameters: 1012620

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/1/paa/pct\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	72.7	14	6	US-10-172-425B-33
2	24	72.7	278	5	US-09-675-784A-11175
3	23	69.7	14	6	US-10-172-425B-37
4	23	69.7	115	5	US-09-134-000C-6190
5	23	69.7	115	5	US-09-134-000C-6190
6	23	69.7	134	6	US-10-424-599-184221
7	23	69.7	228	6	US-10-203-536-1
8	23	69.7	266	5	US-09-751-708A-12
9	23	69.7	281	6	US-10-424-599-179861
10	23	69.7	324	7	US-60-440-068-124
11	23	69.7	627	6	US-10-431-652-7705
12	22	66.7	10	5	US-09-150-947F-3
13	22	66.7	10	5	US-09-150-947F-4
14	22	66.7	10	6	US-10-172-425B-3
15	22	66.7	10	6	US-10-172-425B-4
16	22	66.7	12	5	US-09-150-947F-1
17	22	66.7	12	5	US-09-150-947F-2
18	22	66.7	12	6	US-10-172-425B-1
19	22	66.7	12	6	US-10-172-425B-2
20	22	66.7	12	6	US-10-172-425B-52
21	22	66.7	12	7	US-60-458-305-26
22	22	66.7	12	7	US-60-458-305-29
23	22	66.7	13	5	US-09-150-947F-5
24	22	66.7	13	5	US-09-150-947F-6
25	22	66.7	13	5	US-09-150-947F-11
26	22	66.7	13	6	US-10-172-425B-5

27	22	66.7	13	6	US-10-172-425B-6	Sequence 6, Appl1
28	22	66.7	13	6	US-10-172-425B-11	Sequence 11, Appl1
29	22	66.7	14	5	US-09-150-947F-9	Sequence 9, Appl1
30	22	66.7	14	5	US-09-150-947F-10	Sequence 10, Appl1
31	22	66.7	14	6	US-10-172-425B-9	Sequence 9, Appl1
32	22	66.7	14	6	US-10-172-425B-10	Sequence 10, Appl1
33	22	66.7	14	6	US-10-172-425B-13	Sequence 13, Appl1
34	22	66.7	14	6	US-10-172-425B-15	Sequence 15, Appl1
35	22	66.7	14	6	US-10-172-425B-26	Sequence 26, Appl1
36	22	66.7	14	6	US-10-172-425B-27	Sequence 27, Appl1
37	22	66.7	14	6	US-10-172-425B-36	Sequence 36, Appl1
38	22	66.7	14	6	US-10-172-425B-42	Sequence 42, Appl1
39	22	66.7	14	6	US-10-172-425B-44	Sequence 44, Appl1
40	22	66.7	14	6	US-10-172-425B-45	Sequence 45, Appl1
41	22	66.7	14	6	US-10-172-425B-46	Sequence 46, Appl1
42	22	66.7	14	6	US-10-172-425B-47	Sequence 47, Appl1
43	22	66.7	14	6	US-10-172-425B-50	Sequence 50, Appl1
44	22	66.7	16	6	US-10-172-425B-14	Sequence 14, Appl1
45	22	66.7	24	5	US-09-150-947F-7	Sequence 7, Appl1

## ALIGNMENTS

```

RESULT 1
US-10-172-425B-33
; Sequence 33, Application US/10172425B
; GENERAL INFORMATION:
; APPLICANT: Kaempfer, Raymond
; APPLICANT: Arad, Gila
; TITLE OF INVENTION: BROAD SPECTRUM ANTAGONISTS AND VACCINES
; TITLE OF INVENTION: DIRECTED AGAINST PYROGNIC EXOTOXINS
; FILE REFERENCE: A31967-PCT-USA-A-A 066031.0164
; CURRENT APPLICATION NUMBER: US/10/172.425B
; CURRENT FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 09/150,947
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: PCT/IL97/00438
; PRIOR FILING DATE: 1997-12-30
; PRIOR APPLICATION NUMBER: ISRAEL 119938
; PRIOR FILING DATE: 1996-12-30
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-172-425B-33
Query Match. 72.7%; Score 24; DB 6; Length 14;
Best Local Similarity 50.0%; Pred. No. 4.9;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
CY 3 XXXXXXTXQEXD 12
| | | | |
DB 5 KTVTAQED 14
;
RESULT 2
US-09-675-784A-11175
; Sequence 11175, Application US/09675784A
; GENERAL INFORMATION:
; APPLICANT: HARE, ROBERTA S.
; APPLICANT: SHAW, KAREN J.
; APPLICANT: SHIMER JR., GEORGE H.
; APPLICANT: KESSLER, MARCO
; APPLICANT: NOLLING, JORR
; APPLICANT: ZENG, QIANJONG
; APPLICANT: GREENE, JONATHAN R.
; TITLE OF INVENTION: ASPERGILLUS FUMIGATUS NUCLEIC ACIDS AND POLYPEPTIDES,
; FILE REFERENCE: 2976-4020US1
; CURRENT APPLICATION NUMBER: US/09/675,784A

```

;; CURRENT FILING DATE: 2000-09-29  
;; PRIOR APPLICATION NUMBER: 60/156,338  
;; PRIOR FILING DATE: 1999-09-29  
;; NUMBER OF SEQ ID NOS: 13925  
;; SEQ ID NO 11175  
;; LENGTH: 278  
;; TYPE: PRT  
;; ORGANISM: Aspergillus fumigatus  
US-09-675-784A-11175

Query Match 72.7%; Score 24; DB 5; Length 278;  
Best Local Similarity 50.0%; Pred. No. 85;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
| | | | |  
DB 126 KSDTTQDESD 135

RESULT 3  
US-10-172-425B-37  
; Sequence 37, Application US/10172425B  
; GENERAL INFORMATION:  
; APPLICANT: Kaempfer, Raymond  
; APPLICANT: Arad, Gila  
; TITLE OF INVENTION: BROAD SPECTRUM ANTAGONISTS AND VACCINES  
; TITLE OF INVENTION: DIRECTED AGAINST PYROGENIC EXOTOXINS  
; FILE REFERENCE: A31967-PCT-USA-A-A 066031, 0164  
; CURRENT APPLICATION NUMBER: US/10/172,425B  
; CURRENT FILING DATE: 2002-06-13  
; PRIOR APPLICATION NUMBER: 09/150,947  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: PCT/IL97/00438  
; PRIOR FILING DATE: 1997-12-30  
; PRIOR APPLICATION NUMBER: ISRAEL 119938  
; PRIOR FILING DATE: 1996-12-30  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 37  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-10-172-425B-37

Query Match 69.7%; Score 23; DB 6; Length 14;  
Best Local Similarity 50.0%; Pred. No. 9.2;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
| | | | |  
DB 5 KRSVTAQELD 14

RESULT 4  
US-09-134-000C-6190  
; Sequence 6190, Application US/09134000C  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6190  
; LENGTH: 115  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-6190

Query Match 69.7%; Score 23; DB 5; Length 115;  
Best Local Similarity 50.0%; Pred. No. 69;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
| | | | |  
DB 85 KKESTQSEVD 94

RESULT 5  
US-09-134-000C-6190  
; Sequence 6190, Application US/09134000C  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6190  
; LENGTH: 115  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-6190

Query Match 69.7%; Score 23; DB 5; Length 115;  
Best Local Similarity 50.0%; Pred. No. 69;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
| | | | |  
DB 85 KKESTQSEVD 94

RESULT 6  
US-10-424-599-184221  
; Sequence 184221, Application US/10424599  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 184221  
; LENGTH: 134  
; TYPE: PRT  
; ORGANISM: Glycine max  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_137368C.1 pep  
US-10-424-599-184221

Query Match 69.7%; Score 23; DB 6; Length 134;  
Best Local Similarity 50.0%; Pred. No. 80;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 KXXXTXQEXD 12  
| | | | |  
DB 21 KTHDTSQELD 30

RESULT 7  
US-10-203-536-1  
; Sequence 1, Application US/10203536  
; GENERAL INFORMATION:

APPLICANT: LEE, Hong-Kyun  
APPLICANT: PARK, Yong-Ho  
APPLICANT: HAN, Kyu-Boem  
APPLICANT: CHANG, Byoung-Sun  
APPLICANT: LEE, Yong-Jun  
TITLE OF INVENTION: Staphylococcal Enterotoxin SEC-SER, Expression Vector and Host Cell  
FILE REFERENCE: 0808-0329P  
CURRENT APPLICATION NUMBER: US/10/203,536  
CURRENT FILING DATE: 2002-08-12  
PRIOR APPLICATION NUMBER: PCT/KR00/01241  
PRIOR FILING DATE: 2000-10-31  
PRIOR APPLICATION NUMBER: KR 2000-0007612  
PRIOR FILING DATE: 2000-02-17  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 228  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Staphylococcal enterotoxin SEC-SER  
US-10-203-536-1

Query Match 69.7%; Score 23; DB 6; Length 228;  
Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQXND 12  
DB 140 KKSXTAQEED 149

RESULT 8  
US-09-751-708A-12  
Sequence 12, Application US/09751708A  
GENERAL INFORMATION:  
APPLICANT: TERMAN, David S  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE  
FILE REFERENCE: 751708  
CURRENT APPLICATION NUMBER: US/09/751,708A  
CURRENT FILING DATE: 2002-10-15  
PRIOR APPLICATION NUMBER: US 60/173,371  
PRIOR FILING DATE: 1999-12-28  
NUMBER OF SEQ ID NOS: 166  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 12  
LENGTH: 266  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-751-708A-12

Query Match 69.7%; Score 23; DB 5; Length 266;  
Best Local Similarity 50.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQXND 12  
DB 178 KKSXTAQEED 187

RESULT 9  
US-10-424-599-179861  
Sequence 179861, Application US/10424599  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599

CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 179861  
LENGTH: 281  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_13342C.1.pep  
US-10-424-599-179861

Query Match 69.7%; Score 23; DB 6; Length 281;  
Best Local Similarity 50.0%; Pred. No. 1.6e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQXND 12  
DB 176 KKSXTAQEED 185

RESULT 10  
US-60-440-068-124  
Sequence 124, Application US/60440068  
GENERAL INFORMATION:  
APPLICANT: NADLER, STEVEN G.  
APPLICANT: CARMAN, JULIE  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE  
FILE REFERENCE: 3053-4191  
CURRENT APPLICATION NUMBER: US/60/440,068  
CURRENT FILING DATE: 2003-01-14  
NUMBER OF SEQ ID NOS: 746  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 124  
LENGTH: 324  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-60-440-068-124

Query Match 69.7%; Score 23; DB 7; Length 324;  
Best Local Similarity 50.0%; Pred. No. 1.8e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQXND 12  
DB 193 KKSXTAQEED 202

RESULT 11  
US-10-431-652-7705  
Sequence 7705, Application US/10431652  
GENERAL INFORMATION:  
APPLICANT: Breton, Gary L.  
APPLICANT: Bush, David  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: PAT03-08  
CURRENT APPLICATION NUMBER: US/10/431,652  
CURRENT FILING DATE: 2003-05-06  
PRIOR APPLICATION NUMBER: US 09/328,352  
PRIOR FILING DATE: 1999-06-04  
PRIOR APPLICATION NUMBER: US 60/088,701  
PRIOR FILING DATE: 1998-06-09  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 7705  
LENGTH: 627  
TYPE: PRT  
ORGANISM: Acinetobacter baumannii  
US-10-431-652-7705

Query Match 69.7%; Score 23; DB 6; Length 627;  
Best Local Similarity 50.0%; Pred. No. 3.3e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12  
| | | | |  
Db 355 KSKRTAQELD 364

RESULT 12  
US-09-150-947F-3  
; Sequence 3, Application US/09150947F  
; GENERAL INFORMATION:  
; APPLICANT: Kaempfer, Raymond  
; APPLICANT: Arad, Gila  
; TITLE OF INVENTION: BROAD SPECTRUM PYROGENIC EXOTOXINS ANTAGONISTS VACCINES  
; FILE REFERENCE: A31967-PCT-USA-A 066031.0102  
; CURRENT FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: PCT/IL97/00438  
; PRIOR FILING DATE: 1997-12-30  
; PRIOR APPLICATION NUMBER: ISRAEL 119938  
; PRIOR FILING DATE: 1996-12-30  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-09-150-947F-3

Query Match 66.7%; Score 22; DB 5; Length 10;  
Best Local Similarity 50.0%; Pred. No. 13;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12  
| | | | |  
Db 1 KKKATVQELD 10

RESULT 13  
US-09-150-947F-4  
; Sequence 4, Application US/09150947F  
; GENERAL INFORMATION:  
; APPLICANT: Kaempfer, Raymond  
; APPLICANT: Arad, Gila  
; TITLE OF INVENTION: BROAD SPECTRUM PYROGENIC EXOTOXINS ANTAGONISTS VACCINES  
; FILE REFERENCE: A31967-PCT-USA-A 066031.0102  
; CURRENT FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: PCT/IL97/00438  
; PRIOR FILING DATE: 1997-12-30  
; PRIOR APPLICATION NUMBER: ISRAEL 119938  
; PRIOR FILING DATE: 1996-12-30  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-09-150-947F-4

Query Match 66.7%; Score 22; DB 5; Length 10;  
Best Local Similarity 50.0%; Pred. No. 13;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12  
| | | | |  
Db 1 KKKATVQELD 10

RESULT 14  
US-10-172-425B-3  
; Sequence 3, Application US/10172425B  
; GENERAL INFORMATION:  
; APPLICANT: Kaempfer, Raymond  
; APPLICANT: Arad, Gila

; TITLE OF INVENTION: BROAD SPECTRUM ANTAGONISTS AND VACCINES  
; FILE REFERENCE: DIRECTED AGAINST PYROGENIC EXOTOXINS  
; FILE REFERENCE: A31967-PCT-USA-A 066031.0164  
; CURRENT FILING DATE: 2002-06-13  
; PRIOR APPLICATION NUMBER: 09/150,947  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: PCT/IL97/00438  
; PRIOR FILING DATE: 1997-12-30  
; PRIOR APPLICATION NUMBER: ISRAEL 119938  
; PRIOR FILING DATE: 1996-12-30  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-10-172-425B-3

Query Match 66.7%; Score 22; DB 6; Length 10;  
Best Local Similarity 50.0%; Pred. No. 13;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12  
| | | | |  
Db 1 KKKATVQELD 10

RESULT 15  
US-10-172-425B-4  
; Sequence 4, Application US/10172425B  
; GENERAL INFORMATION:  
; APPLICANT: Kaempfer, Raymond  
; APPLICANT: Arad, Gila  
; TITLE OF INVENTION: BROAD SPECTRUM ANTAGONISTS AND VACCINES  
; FILE REFERENCE: DIRECTED AGAINST PYROGENIC EXOTOXINS  
; FILE REFERENCE: A31967-PCT-USA-A 066031.0164  
; CURRENT FILING DATE: 2002-06-13  
; PRIOR APPLICATION NUMBER: US/10/172,425B  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 09/150,947  
; PRIOR FILING DATE: 1997-12-30  
; PRIOR APPLICATION NUMBER: ISRAEL 119938  
; PRIOR FILING DATE: 1996-12-30  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-10-172-425B-4

Query Match 66.7%; Score 22; DB 6; Length 10;  
Best Local Similarity 50.0%; Pred. No. 13;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 KXXXTXQEXD 12  
| | | | |  
Db 1 KKKATVQELD 10

Search completed: June 5, 2003, 15:52:27  
Job time : 34.1818 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:29:17 ; Search time 147.818 Seconds  
(without alignments)  
52.340 Million cell updates/sec

Title: US-09-150-947f-17  
Perfect score: 28  
Sequence: 1 XXXXXXXXXXXXLD 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database :

Pending\_Patents\_AA\_Main:\*  
1: /cgn2\_6/ptodata/1/paa/US086.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US086.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US087.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US087.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US088.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US088.COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US088.COMB.pep:\*  
8: /cgn2\_6/ptodata/1/paa/US088.COMB.pep:\*  
9: /cgn2\_6/ptodata/1/paa/US088.COMB.pep:\*  
10: /cgn2\_6/ptodata/1/paa/US088.COMB.pep:\*  
11: /cgn2\_6/ptodata/1/paa/US088.COMB.pep:\*  
12: /cgn2\_6/ptodata/1/paa/US088.COMB.pep:\*  
13: /cgn2\_6/ptodata/1/paa/US089.COMB.pep:\*  
14: /cgn2\_6/ptodata/1/paa/US089.COMB.pep:\*  
15: /cgn2\_6/ptodata/1/paa/US091.COMB.pep:\*  
16: /cgn2\_6/ptodata/1/paa/US092.COMB.pep:\*  
17: /cgn2\_6/ptodata/1/paa/US093.COMB.pep:\*  
18: /cgn2\_6/ptodata/1/paa/US094.COMB.pep:\*  
19: /cgn2\_6/ptodata/1/paa/US095.COMB.pep:\*  
20: /cgn2\_6/ptodata/1/paa/US096.COMB.pep:\*  
21: /cgn2\_6/ptodata/1/paa/US097.COMB.pep:\*  
22: /cgn2\_6/ptodata/1/paa/US098.COMB.pep:\*  
23: /cgn2\_6/ptodata/1/paa/US099.COMB.pep:\*  
24: /cgn2\_6/ptodata/1/paa/US100.COMB.pep:\*  
25: /cgn2\_6/ptodata/1/paa/US101.COMB.pep:\*  
26: /cgn2\_6/ptodata/1/paa/US102.COMB.pep:\*  
27: /cgn2\_6/ptodata/1/paa/US60.COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	67.9	310	27	US-60-360-039-12686
2	19	67.9	358	21	US-09-733-089-16924
3	19	67.9	358	22	US-09-816-660-16924
4	19	67.9	378	19	US-09-513-996A-43459
5	19	67.9	378	21	US-09-708-427-18784
6	19	67.9	378	21	US-09-791-537-114720

7	19	67.9	378	27	US-60-356-051-2815	Sequence 2815, Ap
8	18	64.3	35	26	US-10-221-279-10068	Sequence 10068, A
9	18	64.3	43	1	PCT-US99-22853B-42	Sequence 42, Appl
10	18	64.3	46	1	PCT-US00-07527-111	Sequence 111, Appl
11	18	64.3	46	1	PCT-US00-07527-112	Sequence 112, Appl
12	18	64.3	54	1	PCT-US99-22853B-41	Sequence 41, Appl
13	18	64.3	79	1	PCT-US99-22853B-40	Sequence 40, Appl
14	18	64.3	87	15	US-09-134-000-5378	Sequence 5378, Ap
15	18	64.3	89	21	US-09-758-472-8508	Sequence 8508, Ap
16	18	64.3	89	26	US-10-235-926-8508	Sequence 8508, Ap
17	18	64.3	100	19	US-09-565-177A-13	Sequence 13, Appl
18	18	64.3	100	21	US-09-732-210-1170	Sequence 1170, Ap
19	18	64.3	100	21	US-09-791-537-23478	Sequence 23478, A
20	18	64.3	122	20	US-60-169-340-1170	Sequence 1170, Ap
21	18	64.3	122	20	US-09-617-681A-4242	Sequence 4242, Ap
22	18	64.3	131	21	US-09-708-427-46061	Sequence 46061, A
23	18	64.3	140	16	US-09-248-796-25982	Sequence 25982, A
24	18	64.3	140	27	US-60-096-409-25982	Sequence 25982, A
25	18	64.3	141	20	US-09-617-681A-4241	Sequence 4241, Ap
26	18	64.3	152	19	US-09-565-177A-4	Sequence 4, Appl
27	18	64.3	152	19	US-09-565-177A-6	Sequence 6, Appl
28	18	64.3	152	19	US-09-565-177A-12	Sequence 12, Appl
29	18	64.3	160	21	US-09-708-427-20661	Sequence 20661, A
30	18	64.3	171	21	US-09-708-427-20660	Sequence 20660, A
31	18	64.3	173	1	PCT-US01-08631-50628	Sequence 50628, A
32	18	64.3	200	22	US-09-809-391-495	Sequence 495, Appl
33	18	64.3	200	22	US-09-882-171-495	Sequence 495, Appl
34	18	64.3	200	25	US-10-164-861-495	Sequence 495, Appl
35	18	64.3	205	27	US-60-389-987-655	Sequence 655, Appl
36	18	64.3	205	27	US-60-412-418-655	Sequence 655, Appl
37	18	64.3	206	27	US-60-389-987-2179	Sequence 2179, Ap
38	18	64.3	206	27	US-60-412-418-2179	Sequence 2179, Ap
39	18	64.3	237	1	PCT-US02-23913-91	Sequence 91, Appl
40	18	64.3	237	13	US-08-933-750-32	Sequence 32, Appl
41	18	64.3	237	19	US-09-518-865-32	Sequence 32, Appl
42	18	64.3	237	12	US-09-840-787-32	Sequence 32, Appl
43	18	64.3	237	26	US-10-503-823-91	Sequence 91, Appl
44	18	64.3	247	8	US-08-467-153-18	Sequence 18, Appl

## ALIGNMENTS

RESULT 1  
US-60-360-039-12686  
: Sequence 12686, Application US/60360039  
: GENERAL INFORMATION:  
: APPLICANT: Cao, Yongwei  
: APPLICANT: Chen, Xianfeng  
: APPLICANT: Goldman, Barry S.  
: APPLICANT: Hinkle, Gregory J.  
: APPLICANT: Slater, Steven C.  
: TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
: FILE REFERENCE: 38-10(52052)A  
: CURRENT APPLICATION NUMBER: US/60/360,039  
: CURRENT FILING DATE: 2002-02-21  
: NUMBER OF SEQ ID NOS: 47374  
: SEQ ID NO 12686  
: LENGTH: 310  
: TYPE: PRT  
: ORGANISM: Aspergillus nidulans  
: FEATURE:  
: NAME/KEY: unsure  
: LOCATION: (1)..(310)  
: OTHER INFORMATION: unsure at all Xaa locations  
US-60-360-039-12686

Query Match 67.9%; Score 19; DB 27; Length 310;  
Best Local Similarity 40.0%; Pred. No. 2.4e+03;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KXXXXXXXXD 12  
DB 73 KKTAAAFALD 82

## RESULT 2

US-09-733-089-16924  
Sequence 16924, Application US/09733089  
GENERAL INFORMATION:  
APPLICANT: Dotson, Stanton B.  
APPLICANT: Kovacic, David K.  
APPLICANT: Liu, Jingdong  
APPLICANT: Lutfiyya, Linda L.  
APPLICANT: McIninch, James  
APPLICANT: Wu, Wei  
TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With  
FILE REFERENCE: 38-21(15300)D  
CURRENT APPLICATION NUMBER: US/09/733,089  
CURRENT FILING DATE: 2000-12-11  
PRIOR APPLICATION NUMBER: US 09/474,435  
PRIOR FILING DATE: 1999-12-28  
PRIOR APPLICATION NUMBER: US 09/654,617  
PRIOR FILING DATE: 2000-09-05  
PRIOR APPLICATION NUMBER: US 09/620,392  
PRIOR FILING DATE: 2000-07-19  
NUMBER OF SEQ ID NOS: 24143  
SEQ ID NO 16924  
LENGTH: 358  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-09-733-089-16924

Query Match 67.9%; Score 19; DB 21; Length 358;  
Best Local Similarity 40.0%; Pred. No. 2.8e+03;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KXXXXXXXXD 12  
DB 42 KKAADAAALD 51

## RESULT 3

US-09-816-660-16924  
Sequence 16924, Application US/09816660  
GENERAL INFORMATION:  
APPLICANT: Dotson, Stanton B.  
APPLICANT: Kovacic, David K.  
APPLICANT: Liu, Jingdong  
APPLICANT: Lutfiyya, Linda L.  
APPLICANT: McIninch, James  
APPLICANT: Wu, Wei  
TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With  
FILE REFERENCE: 38-21(15300)D  
CURRENT APPLICATION NUMBER: US/09/816,660  
CURRENT FILING DATE: 2001-03-26  
PRIOR APPLICATION NUMBER: US 09/474,435  
PRIOR FILING DATE: 1999-12-28  
PRIOR APPLICATION NUMBER: US 09/654,617  
PRIOR FILING DATE: 2000-09-05  
PRIOR APPLICATION NUMBER: US 09/733,089  
PRIOR FILING DATE: 2000-12-11  
PRIOR APPLICATION NUMBER: ) US 09/684,016  
PRIOR FILING DATE: 2000-10-10  
PRIOR APPLICATION NUMBER: US 09/620,392  
PRIOR FILING DATE: 2000-07-19  
NUMBER OF SEQ ID NOS: 24143  
SEQ ID NO 16924  
LENGTH: 358  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-09-816-660-16924

Query Match 67.9%; Score 19; DB 22; Length 358;  
Best Local Similarity 40.0%; Pred. No. 2.8e+03;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KXXXXXXXXD 12  
DB 42 KKAADAAALD 51

## RESULT 4

US-09-513-996A-43459  
Sequence 43459, Application US/09513996A  
GENERAL INFORMATION:  
APPLICANT: N. ALEXANDROV et al.  
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
FILE REFERENCE: 2750-709P  
CURRENT APPLICATION NUMBER: US/09/513,996A  
CURRENT FILING DATE: 2000-02-25  
NUMBER OF SEQ ID NOS: 81028  
SEQ ID NO 43459  
LENGTH: 378  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: 1..378  
OTHER INFORMATION: any n or xaa - unknown  
FEATURE:  
OTHER INFORMATION: Location 1..378 / Ceres Seq. ID 1985639  
US-09-513-996A-43459

Query Match 67.9%; Score 19; DB 19; Length 378;  
Best Local Similarity 40.0%; Pred. No. 3e+03;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KXXXXXXXXD 12  
DB 13 KKAADAAALD 22

## RESULT 5

US-09-708-427-18784  
Sequence 18784, Application US/09708427  
GENERAL INFORMATION:  
APPLICANT: N. ALEXANDROV et al.  
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
FILE REFERENCE: 2750-1243P  
CURRENT APPLICATION NUMBER: US/09/708,427  
CURRENT FILING DATE: 2000-11-09  
NUMBER OF SEQ ID NOS: 85364  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 18784  
LENGTH: 378  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1..378  
OTHER INFORMATION: xaa is any amino acid.  
NAME/KEY: misc\_feature  
LOCATION: 1..378  
OTHER INFORMATION: Ceres Seq. ID 1834815  
US-09-708-427-18784

Query Match 67.9%; Score 19; DB 21; Length 378;  
Best Local Similarity 40.0%; Pred. No. 3e+03;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KXXXXXXXXD 12  
DB 11

Db 13 KKAADAAALD 22

RESULT 6

US-09-791-537-114720  
; Sequence 114720, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Blonoximix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 114720  
; LENGTH: 378  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-791-537-114720

Query Match  
Best Local Similarity 40.0%; Score 19; DB 21; Length 378;  
Pred. No. 3e+03;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KKKXXXXXLD 12  
||  
Db 13 KKAADAAALD 22

RESULT 7

US-60-356-051-2815  
; Sequence 2815, Application US/60356051  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Technology LLC  
; APPLICANT: Lufflyya, Linda L.  
; TITLE OF INVENTION: NUCLEIC ACIDS AND OTHER MOLECULES ASSOCIATED WITH  
; TITLE OF INVENTION: TRANSCRIPTION IN PLANTS  
; FILE REFERENCE: 38-21 (15300)1  
; CURRENT APPLICATION NUMBER: US/60/356,051  
; CURRENT FILING DATE: 2002-02-11  
; NUMBER OF SEQ ID NOS: 2926  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 2815  
; LENGTH: 378  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-60-356-051-2815

Query Match  
Best Local Similarity 40.0%; Score 19; DB 27; Length 378;  
Pred. No. 3e+03;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KKKXXXXXLD 12  
||  
Db 13 KKAADAAALD 22

RESULT 8

US-10-221-279-10068  
; Sequence 10068, Application US/10221279  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides  
; FILE REFERENCE: 21272-046  
; CURRENT APPLICATION NUMBER: US/10/221,279  
; CURRENT FILING DATE: 2002-09-06  
; PRIOR APPLICATION NUMBER: 09/574,454  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: 09/519,705  
; PRIOR FILING DATE: 2000-03-07

; NUMBER OF SEQ ID NOS: 12360  
; SOFTWARE: Custom  
; SEQ ID NO 10068  
; LENGTH: 35  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-221-279-10068

Query Match  
Best Local Similarity 40.0%; Score 18; DB 26; Length 35;  
Pred. No. 4.5e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KKKXXXXXLD 12  
||  
Db 3 KKYSSQTSLD 12

RESULT 9

PCT-US99-22853B-42  
; Sequence 42, Application PC/TUS9922853B  
; GENERAL INFORMATION:  
; APPLICANT: Ceres, Inc.  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding  
; TITLE OF INVENTION: Polypeptides Encoded Thereby  
; FILE REFERENCE: 2750-0569F(PC)  
; CURRENT APPLICATION NUMBER: PCT/US99/22853B  
; CURRENT FILING DATE: 1999-10-05  
; NUMBER OF SEQ ID NOS: 3938  
; SOFTWARE: MS Word 97  
; SEQ ID NO 42  
; LENGTH: 43  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; OTHER INFORMATION: LOCATION 1..43, Ceres Seq. ID 1685642  
; NAME/KEY: UNSURE  
; LOCATION: (1)..(43)  
; OTHER INFORMATION: any Xaa - any amino acid, unknown or other  
PCT-US99-22853B-42

Query Match  
Best Local Similarity 40.0%; Score 18; DB 1; Length 43;  
Pred. No. 5.7e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KKKXXXXXLD 12  
||  
Db 5 KKLTTDLD 14

RESULT 10

PCT-US00-07527-111  
; Sequence 111, Application PC/TUS0007527  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: 47 Human Secreted Proteins  
; FILE REFERENCE: P5529PCT  
; CURRENT APPLICATION NUMBER: PCT/US00/07527  
; CURRENT FILING DATE: 2000-03-22  
; EARLIER APPLICATION NUMBER: 60/126,600  
; EARLIER FILING DATE: 1999-03-26  
; EARLIER APPLICATION NUMBER: 60/171,550  
; EARLIER FILING DATE: 1999-171,550  
; NUMBER OF SEQ ID NOS: 138  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 111  
; LENGTH: 46  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US00-07527-111

Query Match  
Best Local Similarity 40.0%; Score 18; DB 1; Length 46;  
Pred. No. 6.1e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KXXXXXXXXD 12  
DB 19 KKSAAESMD 28

## RESULT 11

PCT-US00-07527-112  
; Sequence 112, Application PC/TUS0007527  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: 47 Human Secreted Proteins  
; FILE REFERENCE: P5529PCT  
; CURRENT APPLICATION NUMBER: PCT/US00/07527  
; CURRENT FILING DATE: 2000-03-22  
; EARLIER APPLICATION NUMBER: 60/126,600  
; EARLIER FILING DATE: 1999-03-26  
; EARLIER APPLICATION NUMBER: 60/171,550  
; EARLIER FILING DATE: 1999-171,550  
; NUMBER OF SEQ ID NOS: 138  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 112  
; LENGTH: 46  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US00-07527-112

Query Match 64.3%; Score 18; DB 1; Length 46;  
Best Local Similarity 40.0%; Pred. No. 6.1e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KXXXXXXXXD 12  
DB 19 KKSAAESMD 28

## RESULT 12

PCT-US99-22853B-41  
; Sequence 41, Application PC/TUS9922853B  
; GENERAL INFORMATION:  
; APPLICANT: Ceres, Inc.  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding  
; FILE REFERENCE: 2750-0569F(PC)  
; CURRENT APPLICATION NUMBER: PCT/US99/22853B  
; CURRENT FILING DATE: 1999-10-05  
; NUMBER OF SEQ ID NOS: 3938  
; SOFTWARE: MS Word 97  
; SEQ ID NO 41  
; LENGTH: 54  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; OTHER INFORMATION: LOCATION 1..54, Ceres Seq. ID 1685641  
; LOCATION: (1)..(54)  
; OTHER INFORMATION: any xaa - any amino acid, unknown or other  
PCT-US99-22853B-41

Query Match 64.3%; Score 18; DB 1; Length 54;  
Best Local Similarity 40.0%; Pred. No. 7.3e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KXXXXXXXXD 12  
DB 16 KKLTTTDLAD 25

RESULT 13  
PCT-US99-22853B-40  
; Sequence 40, Application PC/TUS9922853B  
; GENERAL INFORMATION:  
; APPLICANT: Ceres, Inc.

; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding  
; FILE REFERENCE: 2750-0569F(PC)  
; CURRENT APPLICATION NUMBER: PCT/US99/22853B  
; CURRENT FILING DATE: 1999-10-05  
; NUMBER OF SEQ ID NOS: 3938  
; SOFTWARE: MS Word 97  
; SEQ ID NO 40  
; LENGTH: 79  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; OTHER INFORMATION: LOCATION 1..79, Ceres Seq. ID 1685640  
; NAME/KEY: UNSURE  
; LOCATION: (1)..(79)  
; OTHER INFORMATION: any xaa - any amino acid, unknown or other  
PCT-US99-22853B-40

Query Match 64.3%; Score 18; DB 1; Length 79;  
Best Local Similarity 40.0%; Pred. No. 1.1e+03;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KXXXXXXXXD 12  
DB 41 KKLTTTDLAD 50

## RESULT 14

US-09-134-000-5378  
; Sequence 5378, Application US/09134000A  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS  
; FILE REFERENCE: GTC-005  
; CURRENT APPLICATION NUMBER: US/09/134,000A  
; CURRENT FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 6810  
; SEQ ID NO 5378  
; LENGTH: 87  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-134-000-5378

Query Match 64.3%; Score 18; DB 15; Length 87;  
Best Local Similarity 40.0%; Pred. No. 1.2e+03;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KXXXXXXXXD 12  
DB 44 KKSRAFASLD 53

## RESULT 15

US-09-758-472-8508  
; Sequence 8508, Application US/09758472  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PH001  
; CURRENT APPLICATION NUMBER: US/09/758,472  
; CURRENT FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; NUMBER OF SEQ ID NOS: 9632  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 8508  
; LENGTH: 89  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:

US-09-758-472-8508



```

; NAME/KEY: SITE
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (4)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (18)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-758-472-8508

```

```

Query Match      64.3%; Score 18; DB 21; Length 89;
Best Local Similarity 40.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

QY      3 KKKXXXXXD 12
      ||      ||
Db      43 KKAIIAALD 52

```

Search completed: June 5, 2003, 15:47:37  
 Job time : 147.818 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 15:31:47 ; Search time 34.1818 Seconds  
(without alignments)  
74.544 Million cell updates/sec

Title: US-09-150-947f-17  
Perfect score: 28  
Sequence: 1 XXXXXXXXXXXXLD 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1012620 segs, 21337834 residues

Total number of hits satisfying chosen parameters: 1012620

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*  
1: /cgn2\_6/pcdata/1/paa/PCOT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/pcdata/1/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/pcdata/1/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/pcdata/1/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/pcdata/1/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/pcdata/1/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/pcdata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67.9	310	6	US-10-369-493-12686	Sequence 12686, A
2	67.9	378	6	US-10-361-942-986	Sequence 986, App
3	64.3	87	5	US-09-134-000C-5378	Sequence 5378, Ap
4	64.3	87	5	US-09-134-000C-5378	Sequence 5378, Ap
5	64.3	100	6	US-10-421-684-1170	Sequence 1170, Ap
6	64.3	109	5	US-09-724-676-55061	Sequence 55061, A
7	64.3	109	5	US-09-724-676-55061	Sequence 55061, A
8	64.3	147	5	US-09-513-999C-6288	Sequence 6288, Ap
9	64.3	147	5	US-09-513-999C-6288	Sequence 6288, Ap
10	64.3	152	5	US-09-949-016-6460	Sequence 6460, Ap
11	64.3	152	7	US-06-440-068-114	Sequence 114, App
12	64.3	161	5	US-09-724-676-55060	Sequence 55060, A
13	64.3	161	5	US-09-724-676-55060	Sequence 55060, A
14	64.3	252	6	US-10-282-122A-55060	Sequence 52371, A
15	64.3	315	6	US-10-431-652-7651	Sequence 7651, Ap
16	64.3	315	6	US-10-431-652-7651	Sequence 4290, Ap
17	64.3	318	6	US-10-272-490-90	Sequence 90, App1
18	64.3	318	6	US-10-272-490-90	Sequence 92, App1
19	64.3	328	6	US-10-363-616-281	Sequence 281, App1
20	64.3	341	6	US-10-272-490-68	Sequence 68, App1
21	64.3	341	6	US-10-272-490-78	Sequence 78, App1
22	64.3	341	6	US-10-214-473-68	Sequence 68, App1
23	64.3	341	6	US-10-214-473-68	Sequence 78, App1
24	64.3	406	6	US-10-425-114-70484	Sequence 70484, A
25	64.3	768	6	US-10-218-140-4424	Sequence 4424, Ap
26	64.3	837	6	US-10-369-493-12296	Sequence 22296, A

ALIGNMENTS

27	17	60.7	12	6	US-10-172-425B-53	Sequence 53, App1
28	17	60.7	14	6	US-10-172-425B-37	Sequence 37, App1
29	17	60.7	14	6	US-10-172-425B-49	Sequence 49, App1
30	17	60.7	53	6	US-10-424-599-172320	Sequence 172320, App1
31	17	60.7	57	6	US-10-424-599-172320	Sequence 172320, App1
32	17	60.7	70	6	US-10-424-599-23335	Sequence 23335, App1
33	17	60.7	77	6	US-10-424-599-23335	Sequence 23335, App1
34	17	60.7	85	6	US-10-424-599-23335	Sequence 23335, App1
35	17	60.7	106	5	US-09-864-408A-3042	Sequence 194596, App1
36	17	60.7	109	6	US-10-424-599-167306	Sequence 167306, App1
37	17	60.7	128	6	US-10-357-886-44	Sequence 44, App1
38	17	60.7	142	6	US-09-857-826B-18	Sequence 18, App1
39	17	60.7	143	6	US-10-276-774-2588	Sequence 2588, App1
40	17	60.7	152	6	US-10-417-884-3744	Sequence 3744, App1
41	17	60.7	160	6	US-10-424-599-170060	Sequence 170060, App1
42	17	60.7	170	6	US-10-264-237-2676	Sequence 2676, App1
43	17	60.7	178	6	US-10-282-122A-44631	Sequence 44631, App1
44	17	60.7	189	6	US-10-094-749-2347	Sequence 2347, App1
45	17	60.7	192	5	US-09-675-784A-8051	Sequence 8051, App1

RESULT 1  
US-10-369-493-12686  
Sequence 12686, Application US/10369493  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
PRIOR FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 12686  
LENGTH: 310  
TYPE: PRT  
ORGANISM: Aspergillus nidulans  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(310)  
OTHER INFORMATION: unsure at all Xaa locations  
US-10-369-493-12686  
Query Match 67.9% Score 19; DB 6; Length 310;  
Best Local Similarity 40.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
Cy 3 KXXXXXXXXLD 12  
Db 73 KKTAAAFALD 82  
RESULT 2  
US-10-361-942-986  
Sequence 986, Application US/10361942  
GENERAL INFORMATION:  
APPLICANT: Lutiya, Linda L.  
TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated with  
FILE REFERENCE: 38-21(15300)K  
CURRENT APPLICATION NUMBER: US/10/361,942  
PRIOR FILING DATE: 2003-02-10  
PRIOR APPLICATION NUMBER: US 60/356,051  
PRIOR FILING DATE: 2002-02-11  
NUMBER OF SEQ ID NOS: 2906

;; SEQ ID NO 986  
;; LENGTH: 378  
;; TYPE: PRT  
;; ORGANISM: Arabidopsis thaliana  
;; FEATURE:  
;; OTHER INFORMATION: Clone ID: AT\_T6A23.C1.p24.tg  
US-10-361-942-986

Query Match 67.9%; Score 19; DB 6; Length 378;  
Best Local Similarity 40.0%; Pred. No. 2.1e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KKKXXXXXLD 12  
DB 13 KKAADAAALD 22

RESULT 3  
US-09-134-000C-5378  
; Sequence 5378, Application US/09134000C

;; GENERAL INFORMATION:  
;; APPLICANT: Lynn Doucette-Stamm et al  
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
;; FILE REFERENCE: 032796-032  
;; CURRENT FILING DATE: 1998-08-13  
;; PRIOR APPLICATION NUMBER: US 60/055,778  
;; PRIOR FILING DATE: 1997-08-15  
;; NUMBER OF SEQ ID NOS: 6812  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 5378  
;; LENGTH: 87  
;; TYPE: PRT  
;; ORGANISM: Enterococcus faecalis  
US-09-134-000C-5378

Query Match 64.3%; Score 18; DB 5; Length 87;  
Best Local Similarity 40.0%; Pred. No. 1e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KKKXXXXXLD 12  
DB 44 KKSRAFASLD 53

RESULT 4  
US-09-134-000C-5378  
; Sequence 5378, Application US/09134000C

;; GENERAL INFORMATION:  
;; APPLICANT: Lynn Doucette-Stamm et al  
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
;; FILE REFERENCE: 032796-032  
;; CURRENT FILING DATE: 1998-08-13  
;; PRIOR APPLICATION NUMBER: US 60/055,778  
;; PRIOR FILING DATE: 1997-08-15  
;; NUMBER OF SEQ ID NOS: 6812  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 5378  
;; LENGTH: 87  
;; TYPE: PRT  
;; ORGANISM: Enterococcus faecalis  
US-09-134-000C-5378

Query Match 64.3%; Score 18; DB 5; Length 87;  
Best Local Similarity 40.0%; Pred. No. 1e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KKKXXXXXLD 12  
DB 44 KKSRAFASLD 53

RESULT 5  
US-10-421-684-1170  
; Sequence 1170, Application US/10421684

;; GENERAL INFORMATION:  
;; APPLICANT: Bunkers, Greg J.  
;; APPLICANT: Liang, Jihong  
;; APPLICANT: Mitanck, Cindy A.  
;; APPLICANT: Seale, Jeffrey W.  
;; APPLICANT: Wu, Yonnie S.  
;; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use  
;; FILE REFERENCE: 38-21(15036)B  
;; CURRENT FILING DATE: 2003-04-23  
;; PRIOR APPLICATION NUMBER: US/09/732,210  
;; PRIOR FILING DATE: 2000-12-07  
;; PRIOR APPLICATION NUMBER: US 60/169,513  
;; PRIOR FILING DATE: 1999-12-07  
;; PRIOR APPLICATION NUMBER: US 60/169,340  
;; NUMBER OF SEQ ID NOS: 1753  
;; SEQ ID NO 1170  
;; LENGTH: 100  
;; TYPE: PRT  
;; ORGANISM: Marchantia polymorpha  
US-10-421-684-1170

Query Match 64.3%; Score 18; DB 6; Length 100;  
Best Local Similarity 40.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KKKXXXXXLD 12  
DB 29 KKITESSLD 38

RESULT 6  
US-09-724-676-55061  
; Sequence 55061, Application US/09724676

;; GENERAL INFORMATION:  
;; APPLICANT: Compugen LTD  
;; TITLE OF INVENTION: Variants of alternative splicing  
;; FILE REFERENCE: 129181.4 Compugen  
;; CURRENT FILING DATE: 2000-11-28  
;; NUMBER OF SEQ ID NOS: 97222  
;; SOFTWARE: PatentIn version 3.2  
;; SEQ ID NO 55061  
;; LENGTH: 109  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-724-676-55061

Query Match 64.3%; Score 18; DB 5; Length 109;  
Best Local Similarity 40.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 KKKXXXXXLD 12  
DB 44 KKSAAESMLD 53

RESULT 7  
US-09-724-676A-55061  
; Sequence 55061, Application US/09724676A

;; GENERAL INFORMATION:  
;; APPLICANT: Compugen LTD  
;; TITLE OF INVENTION: Variants of alternative splicing  
;; FILE REFERENCE: 129181.4 Compugen  
;; CURRENT FILING DATE: 2000-11-28  
;; NUMBER OF SEQ ID NOS: 97222

SOFTWARE: PatentIn version 3.2  
SEQ ID NO 55061  
LENGTH: 109  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-724-676A-55061

Query Match  
Best Local Similarity 64.3%; Score 18; DB 5; Length 109;  
40.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KKKXXXXXID 12  
DB 44 KKSAAESMID 53

RESULT 8  
US-09-513-999C-6288  
Sequence 6288, Application US/09513999C  
GENERAL INFORMATION:  
APPLICANT: Dumas Maline Edwards, J.B.  
APPLICANT: Duclet, A.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
FILE REFERENCE: 59, 052, REG  
CURRENT APPLICATION NUMBER: US/09/513, 999C  
CURRENT FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/122,487  
PRIOR FILING DATE: 1999-02-26  
NUMBER OF SEQ ID NOS: 36681  
SOFTWARE: Patent.pm  
SEQ ID NO 6288  
LENGTH: 147  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-513-999C-6288

Query Match  
Best Local Similarity 64.3%; Score 18; DB 5; Length 147;  
40.0%; Pred. No. 1.8e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KKKXXXXXID 12  
DB 78 KKKRATYID 87

RESULT 9  
US-09-513-999C-6288  
Sequence 6288, Application US/09513999C  
GENERAL INFORMATION:  
APPLICANT: Dumas Maline Edwards, J.B.  
APPLICANT: Duclet, A.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
FILE REFERENCE: 59, 052, REG  
CURRENT APPLICATION NUMBER: US/09/513, 999C  
CURRENT FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/122,487  
PRIOR FILING DATE: 1999-02-26  
NUMBER OF SEQ ID NOS: 36681  
SOFTWARE: Patent.pm  
SEQ ID NO 6288  
LENGTH: 147  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-513-999C-6288

Query Match  
Best Local Similarity 64.3%; Score 18; DB 5; Length 147;  
40.0%; Pred. No. 1.8e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KKKXXXXXID 12  
DB 78 KKKRATYID 87

DB 78 KKKRATYID 87

RESULT 10  
US-09-949-016-6460  
Sequence 6460, Application US/09949016  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: C1001307  
CURRENT APPLICATION NUMBER: US/09/949, 016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6460  
LENGTH: 152  
TYPE: PRT  
ORGANISM: Human  
US-09-949-016-6460

Query Match  
Best Local Similarity 64.3%; Score 18; DB 5; Length 152;  
40.0%; Pred. No. 1.9e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KKKXXXXXID 12  
DB 44 KKSAAESMID 53

RESULT 11  
US-60-440-068-114  
Sequence 114, Application US/60440068  
GENERAL INFORMATION:  
APPLICANT: MADLER, STEVEN G.  
APPLICANT: CARMAN, JULIE  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE  
FILE REFERENCE: 3053-4191  
CURRENT APPLICATION NUMBER: US/60/440,068  
CURRENT FILING DATE: 2003-01-14  
NUMBER OF SEQ ID NOS: 746  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 114  
LENGTH: 152  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-60-440-068-114

Query Match  
Best Local Similarity 64.3%; Score 18; DB 7; Length 152;  
40.0%; Pred. No. 1.9e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KKKXXXXXID 12  
DB 44 KKSAAESMID 53

RESULT 12  
US-09-724-676-55060  
Sequence 55060, Application US/09724676  
GENERAL INFORMATION:  
APPLICANT: Compugen LTD  
TITLE OF INVENTION: Variants of alternative splicing  
FILE REFERENCE: 129181, 4 Compugen  
CURRENT APPLICATION NUMBER: US/09/724,676  
CURRENT FILING DATE: 2000-11-28

NUMBER OF SEQ ID NOS: 97222  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 55060  
LENGTH: 161  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-724-676-55060

Query Match  
Best Local Similarity 64.3%; Score 18; DB 5; Length 161;  
40.0%; Pred. No. 2e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KKKKXXXXLD 12  
11  
11  
DB 44 KKSAAESMLD 53

RESULT 13  
US-09-724-676A-55060  
Sequence 55060, Application US/09724676A  
GENERAL INFORMATION:  
APPLICANT: Compugen LTD  
TITLE OF INVENTION: Variants of alternative splicing  
FILE REFERENCE: 129181.4 Compugen  
CURRENT APPLICATION NUMBER: US/09/724,676A  
CURRENT FILING DATE: 2000-11-28  
NUMBER OF SEQ ID NOS: 97222  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 55060  
LENGTH: 161  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-724-676A-55060

Query Match  
Best Local Similarity 64.3%; Score 18; DB 5; Length 161;  
40.0%; Pred. No. 2e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KKKKXXXXLD 12  
11  
11  
DB 44 KKSAAESMLD 53

RESULT 14  
US-10-282-122A-52371  
Sequence 52371, Application US/10282122A  
GENERAL INFORMATION:  
APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trivick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA 034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See file wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 52371  
LENGTH: 252  
TYPE: PRT  
ORGANISM: Clostridium botulinum  
US-10-282-122A-52371

Query Match  
Best Local Similarity 64.3%; Score 18; DB 6; Length 252;  
40.0%; Pred. No. 3.2e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KKKKXXXXLD 12  
11  
11  
DB 9 KKASTDFTMD 18

RESULT 15  
US-10-431-652-7651  
Sequence 7651, Application US/10431652  
GENERAL INFORMATION:  
APPLICANT: Breton, Gary L.  
APPLICANT: Bush, David  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: PAT03-08  
CURRENT APPLICATION NUMBER: US/10/431,652  
CURRENT FILING DATE: 2003-05-06  
PRIOR APPLICATION NUMBER: US 09/328,352  
PRIOR FILING DATE: 1999-06-04  
PRIOR APPLICATION NUMBER: US 60/088,701  
PRIOR FILING DATE: 1998-06-09  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 7651  
LENGTH: 254  
TYPE: PRT  
ORGANISM: Acinetobacter baumannii  
US-10-431-652-7651

Query Match  
Best Local Similarity 64.3%; Score 18; DB 6; Length 254;  
40.0%; Pred. No. 3.2e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 KKKKXXXXLD 12  
11  
11  
DB 210 KKSAGSKTMD 219

Search completed: June 5, 2003, 15:52:27  
Job time : 34.1818 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: June 5, 2003, 15:29:17 ; Search time 147.818 Seconds  
(without alignments)  
52,340 Million cell updates/sec

Title: US-09-150-947F-18

Perfect score: 17

Sequence: 1 xxxxxxxxxxxx 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US097\_COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US099\_COMB.pep.\*  
24: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*  
25: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*  
26: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*  
27: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35.3	2	1	PCT-US00-00358-13	Sequence 13, Appl
2	35.3	2	1	PCT-US00-08879-16	Sequence 16, Appl
3	35.3	2	1	PCT-US00-08879-18	Sequence 18, Appl
4	35.3	2	1	PCT-US00-08879-30	Sequence 30, Appl
5	35.3	2	1	PCT-US00-08879A-16	Sequence 16, Appl
6	35.3	2	1	PCT-US00-08879A-18	Sequence 18, Appl

7	35.3	2	1	PCT-US00-08879A-30	Sequence 30, Appl
8	35.3	2	1	PCT-US02-09188-532	Sequence 532, App
9	35.3	2	1	PCT-US02-09239-152	Sequence 152, App
10	35.3	2	1	PCT-US02-09370-551	Sequence 551, App
11	35.3	2	1	PCT-US02-23763-296	Sequence 296, App
12	35.3	2	1	PCT-US02-30452-85	Sequence 85, Appl
13	35.3	2	3	PCT-US99-26100-6	Sequence 6, Appl1
14	35.3	2	6	US-07-853-042A-3	Sequence 3, Appl1
15	35.3	2	6	US-08-222-626A-122	Sequence 122, App
16	35.3	2	8	US-08-248-816-8	Sequence 8, Appl1
17	35.3	2	8	US-08-430-121-8	Sequence 8, Appl1
18	35.3	2	8	US-08-472-679B-15	Sequence 15, Appl
19	35.3	2	8	US-08-472-679B-23	Sequence 23, Appl
20	35.3	2	8	US-08-472-679B-33	Sequence 33, Appl
21	35.3	2	8	US-08-472-679B-39	Sequence 39, Appl
22	35.3	2	8	US-08-472-679D-15	Sequence 15, Appl
23	35.3	2	8	US-08-472-679D-23	Sequence 23, Appl
24	35.3	2	8	US-08-472-679D-33	Sequence 33, Appl
25	35.3	2	8	US-08-472-679D-39	Sequence 39, Appl
26	35.3	2	8	US-08-472-679H-15	Sequence 15, Appl
27	35.3	2	8	US-08-472-679H-23	Sequence 23, Appl
28	35.3	2	8	US-08-472-679H-33	Sequence 33, Appl
29	35.3	2	8	US-08-472-679H-39	Sequence 39, Appl
30	35.3	2	10	US-08-633-679B-85	Sequence 85, Appl
31	35.3	2	11	US-08-783-75A-85	Sequence 85, Appl
32	35.3	2	12	US-08-843-675A-72	Sequence 72, Appl
33	35.3	2	16	US-09-295-846-24	Sequence 24, Appl
34	35.3	2	16	US-09-295-846-26	Sequence 26, Appl
35	35.3	2	16	US-09-295-846-38	Sequence 38, Appl
36	35.3	2	16	US-09-295-846A-24	Sequence 24, Appl
37	35.3	2	16	US-09-295-846A-26	Sequence 26, Appl
38	35.3	2	16	US-09-295-846A-38	Sequence 38, Appl
39	35.3	2	16	US-09-295-846B-24	Sequence 24, Appl
40	35.3	2	16	US-09-295-846B-26	Sequence 26, Appl
41	35.3	2	16	US-09-295-846B-38	Sequence 38, Appl
42	35.3	2	16	US-09-295-846B-39	Sequence 39, Appl
43	35.3	2	16	US-09-295-846B-20	Sequence 20, Appl
44	35.3	2	16	US-09-295-846B-32	Sequence 32, Appl
45	35.3	2	16	US-09-296-113A-21	Sequence 21, Appl

## ALIGNMENTS

RESULT 1  
PCT-US00-00358-13  
Sequence 13, Application PC/TUS0000358  
GENERAL INFORMATION:  
APPLICANT: pecker, Irls  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY DISTANTLY HOM  
FILE REFERENCE: 00/20105  
CURRENT APPLICATION NUMBER: PCT/US00/00358  
CURRENT FILING DATE: 2000-01-07  
PRIOR APPLICATION NUMBER: 60/140,801  
PRIOR FILING DATE: 1999-06-25  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 13  
LENGTH: 2  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US00-00358-13

Query Match 35.3%, Score 6, DB 1, Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 12 D 12  
Db 1 D 1

RESULT 2  
PCT-US00-08879-16  
; Sequence 16, Application PC/TUS0008879  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, John  
; APPLICANT: Brandt, Alan  
; APPLICANT: Borovski, Dov  
; TITLE OF INVENTION: Compositions and Methods for Controlling Pests  
; FILE REFERENCE: 4137-120 PCT  
; CURRENT APPLICATION NUMBER: PCT/US00/08879  
; CURRENT FILING DATE: 2000-04-04  
; PRIOR FILING DATE: 1999-04-21  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: TMOF Analogue  
PCT-US00-08879-16

Query Match 35.3%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
|  
Db 1 D 1

RESULT 3  
PCT-US00-08879-18  
; Sequence 18, Application PC/TUS0008879  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, John  
; APPLICANT: Brandt, Alan  
; APPLICANT: Borovski, Dov  
; TITLE OF INVENTION: Compositions and Methods for Controlling Pests  
; FILE REFERENCE: 4137-120 PCT  
; CURRENT APPLICATION NUMBER: PCT/US00/08879  
; CURRENT FILING DATE: 2000-04-04  
; PRIOR APPLICATION NUMBER: 09/295,924  
; PRIOR FILING DATE: 1999-04-21  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: TMOF Analogue  
PCT-US00-08879-18

Query Match 35.3%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
|  
Db 1 D 1

RESULT 4  
PCT-US00-08879-30  
; Sequence 30, Application PC/TUS0008879  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, John  
; APPLICANT: Brandt, Alan  
; APPLICANT: Borovski, Dov  
; TITLE OF INVENTION: Compositions and Methods for Controlling Pests  
; FILE REFERENCE: 4137-120 PCT

; CURRENT APPLICATION NUMBER: PCT/US00/08879  
; CURRENT FILING DATE: 2000-04-04  
; PRIOR APPLICATION NUMBER: 09/295,924  
; PRIOR FILING DATE: 1999-04-21  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 30  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: TMOF Analogue  
PCT-US00-08879-30

Query Match 35.3%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
|  
Db 2 D 2

RESULT 5  
PCT-US00-08879A-16  
; Sequence 16, Application PC/TUS0008879A  
; GENERAL INFORMATION:  
; APPLICANT: Insect Biotechnology, Inc.  
; TITLE OF INVENTION: Compositions and Methods for Controlling Pests  
; FILE REFERENCE: 4137-120 PCT  
; CURRENT APPLICATION NUMBER: PCT/US00/08879A  
; CURRENT FILING DATE: 2000-04-04  
; PRIOR APPLICATION NUMBER: 09/295,924  
; PRIOR FILING DATE: 1999-04-21  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: TMOF Analogue  
PCT-US00-08879A-16

Query Match 35.3%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
|  
Db 1 D 1

RESULT 6  
PCT-US00-08879A-18  
; Sequence 18, Application PC/TUS0008879A  
; GENERAL INFORMATION:  
; APPLICANT: Insect Biotechnology, Inc.  
; TITLE OF INVENTION: Compositions and Methods for Controlling Pests  
; FILE REFERENCE: 4137-120 PCT  
; CURRENT APPLICATION NUMBER: PCT/US00/08879A  
; CURRENT FILING DATE: 2000-04-04  
; PRIOR APPLICATION NUMBER: 09/295,924  
; PRIOR FILING DATE: 1999-04-21  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: TMOF Analogue  
PCT-US00-08879A-18



Query Match 35.3%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
DB 1 D 1

## RESULT 7

PCT-US00-08879A-30  
; Sequence 30, Application PC/TUS0008879A  
; GENERAL INFORMATION:  
; APPLICANT: Insect Biotechnology, Inc.  
; TITLE OF INVENTION: Compositions and Methods for Controlling Pests  
; FILE REFERENCE: 4137-120 PCT  
; CURRENT APPLICATION NUMBER: PCT/US00/08879A  
; PRIOR FILING DATE: 2000-04-04  
; PRIOR FILING DATE: 1999-04-21  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 30  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: TMOF Analogue  
PCT-US00-08879A-30

Query Match 35.3%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
DB 2 D 2

## RESULT 8

PCT-US02-09188-532  
; Sequence 532, Application PC/TUS0209188  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Human Secreted Proteins  
; FILE REFERENCE: PS953PCT  
; CURRENT APPLICATION NUMBER: PCT/US02/09188  
; PRIOR FILING DATE: 2002-03-26  
; PRIOR APPLICATION NUMBER: US 60/278,650  
; PRIOR FILING DATE: 2001-03-27  
; PRIOR APPLICATION NUMBER: US 09/950,082  
; PRIOR FILING DATE: 2001-09-12  
; PRIOR APPLICATION NUMBER: US 09/950,083  
; NUMBER OF SEQ ID NOS: 1732  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 532  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US02-09188-532

Query Match 35.3%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
DB 1 D 1

## RESULT 9

PCT-US02-09239-152  
; Sequence 152, Application PC/TUS0209239  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Human Secreted Proteins  
; FILE REFERENCE: PS953PCT  
; CURRENT APPLICATION NUMBER: PCT/US02/09239  
; PRIOR FILING DATE: 2002-03-26  
; PRIOR APPLICATION NUMBER: US 60/278,650  
; PRIOR FILING DATE: 2001-03-27  
; PRIOR APPLICATION NUMBER: US 09/950,082  
; PRIOR FILING DATE: 2001-09-12  
; PRIOR APPLICATION NUMBER: US 09/950,083  
; NUMBER OF SEQ ID NOS: 380  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 152  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US02-09239-152

Query Match 35.3%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
DB 1 D 1

## RESULT 10

PCT-US02-09370-551  
; Sequence 551, Application PC/TUS0209370  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Human Secreted Proteins  
; FILE REFERENCE: PS954PCT  
; CURRENT APPLICATION NUMBER: PCT/US02/09370  
; PRIOR FILING DATE: 2002-03-26  
; PRIOR APPLICATION NUMBER: US 60/278,650  
; PRIOR FILING DATE: 2001-03-27  
; PRIOR APPLICATION NUMBER: US 09/950,082  
; PRIOR FILING DATE: 2001-09-12  
; PRIOR APPLICATION NUMBER: US 09/950,083  
; NUMBER OF SEQ ID NOS: 1834  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 551  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US02-09370-551

Query Match 35.3%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
DB 1 D 1

## RESULT 11

PCT-US02-23763-296  
; Sequence 296, Application PC/TUS0223763  
; GENERAL INFORMATION:  
; APPLICANT: Sundaramoorthy, M.  
; TITLE OF INVENTION: Crystallized structure of Type IV Collagen NCI Domain Hexamer  
; FILE REFERENCE: MHB 01-1017  
; CURRENT APPLICATION NUMBER: PCT/US02/23763  
; PRIOR FILING DATE: 2002-07-26

PRIOR APPLICATION NUMBER: US 60/308,523  
PRIOR FILING DATE: 2001-07-27  
PRIOR APPLICATION NUMBER: US 60/351,289  
PRIOR FILING DATE: 2001-10-29  
PRIOR APPLICATION NUMBER: US 60/366,854  
PRIOR FILING DATE: 2002-03-22  
PRIOR APPLICATION NUMBER: US 60/385,362  
PRIOR FILING DATE: 2002-06-03  
NUMBER OF SEQ ID NOS: 307  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO: 296  
LENGTH: 2  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US02-23763-296

Query Match 35.3%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12  
DB 2 D 2

RESULT 12  
PCT-US02-30452-85  
Sequence 85, Application PC/TUS0230452  
GENERAL INFORMATION:  
APPLICANT: Wyeth  
TITLE OF INVENTION: Antibody Inhibitors of GDF-8 and Uses Therefor  
FILE REFERENCE: 08/02,00012-00304  
CURRENT APPLICATION NUMBER: PCT/US02/30452  
CURRENT FILING DATE: 2002-09-26  
PRIOR APPLICATION NUMBER: 60/324,528  
PRIOR FILING DATE: 2001-09-26  
NUMBER OF SEQ ID NOS: 131  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO: 85  
LENGTH: 2  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US02-30452-85

Query Match 35.3%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12  
DB 2 D 2

RESULT 13  
PCT-US99-26100-6  
Sequence 6, Application PC/TUS9926100  
GENERAL INFORMATION:  
APPLICANT: Ream, Walt et al.  
TITLE OF INVENTION: Plants Having Enhanced Gall Resistance and  
FILE REFERENCE: 53629  
CURRENT APPLICATION NUMBER: PCT/US99/26100  
CURRENT FILING DATE: 1999-11-04  
EARLIER APPLICATION NUMBER: 60/107,185  
EARLIER FILING DATE: 1998-11-05  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO: 6  
LENGTH: 2  
TYPE: PRT  
ORGANISM: Agrobacterium tumefaciens  
PCT-US99-26100-6

Query Match 35.3%; Score 6; DB 1; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12  
DB 2 D 2

RESULT 14  
US-07-853-042A-3  
Sequence 3, Application US/07853042A  
GENERAL INFORMATION:  
APPLICANT: Anders BIL  
TITLE OF INVENTION: METHOD AND MEANS FOR INDUCING, RESP.,  
TITLE OF INVENTION: PREVENTING CONSTRUCTION OF THE PUPIL  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FRED PHILPITT  
STREET: 99 Canal Center Plaza, Ste. 300  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22314

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: DOS Text File  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/853,042A  
FILING DATE: 19920318  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: FRED PHILPITT  
REGISTRATION NUMBER: 16825  
REFERENCE/DOCKET NUMBER: PHE-808  
TELEPHONE: 703 684 1111  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Polypeptide  
US-07-853-042A-3

Query Match 35.3%; Score 6; DB 3; Length 2;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12  
DB 1 D 1

RESULT 15  
US-08-222-626A-122  
Sequence 122, Application US/08222626A  
GENERAL INFORMATION:  
APPLICANT: Arai, Naoko  
APPLICANT: Masuda, Esteban S.  
APPLICANT: Tokumitsu, Hiroshi  
TITLE OF INVENTION: PURIFIED COMPONENTS OF MAMMALIAN  
TITLE OF INVENTION: TRANSCRIPTION REGULATION COMPLEXES, AND ANALOGS  
NUMBER OF SEQUENCES: 183  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DNA Research Institute  
STREET: 901 California Avenue  
CITY: Palo Alto  
STATE: California

COUNTRY: USA  
 ZIP: 94304-1104  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/222,626A  
 FILING DATE: 04-APR-1994  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/148,061  
 FILING DATE: 05-NOV-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/113,971  
 FILING DATE: 30-AUG-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/088,998  
 FILING DATE: 30-JUL-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/088,483  
 FILING DATE: 06-JUL-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Chang, Edwin P.  
 REGISTRATION NUMBER: 34,090  
 REFERENCE/DOCKET NUMBER: DX0392K3  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-852-9196  
 TELEFAX: 415-496-1200  
 INFORMATION FOR SEQ ID NO: 122:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-222-626A-122

Query Match 35.3%; Score 6; DB 6; Length 2;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 12 D 12  
 Db 1 D 1

Search completed: June 5, 2003, 15:47:38  
 Job time: 148.818 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using SW model

Run on: June 5, 2003, 15:31:47 ; Search time 34.1818 Seconds  
(Without alignments)  
74.544 Million cell updates/sec

Title: US-09-150-947f-18  
Perfect score: 17  
Sequence: 1 xxxxxxxxxxxx 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1012620 seqs, 212337834 residues

Total number of hits satisfying chosen parameters: 1012620

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*  
1: /cgn2\_6/prodata1/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/prodata1/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/prodata1/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/prodata1/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/prodata1/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/prodata1/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/prodata1/paa/US10\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6	35.3	2	4	US-08-472-679H-15
2	6	35.3	2	4	US-08-472-679H-23
3	6	35.3	2	4	US-08-472-679H-33
4	6	35.3	2	4	US-08-472-679H-39
5	6	35.3	2	4	US-08-472-679H-15
6	6	35.3	2	4	US-08-472-679H-23
7	6	35.3	2	4	US-08-472-679H-33
8	6	35.3	2	4	US-08-472-679H-39
9	6	35.3	2	6	US-10-109-048-461
10	6	35.3	2	6	US-10-405-027-5115
11	6	35.3	3	1	PCT-US02-14358-15
12	6	35.3	3	1	PCT-US02-14358-17
13	6	35.3	3	1	PCT-US02-14358-17
14	6	35.3	3	1	PCT-US02-14358-17
15	6	35.3	3	1	PCT-US02-14358-18
16	6	35.3	3	1	PCT-US02-14358-19
17	6	35.3	3	1	PCT-US02-14358-20
18	6	35.3	3	1	PCT-US02-21322-6
19	6	35.3	3	1	PCT-US02-21322A-6
20	6	35.3	3	1	PCT-US02-26918A-316
21	6	35.3	3	1	PCT-US02-26918A-317
22	6	35.3	3	1	PCT-US02-26918A-318
23	6	35.3	3	1	PCT-US02-26918A-319
24	6	35.3	3	1	PCT-US03-00079-30
25	6	35.3	3	1	PCT-US03-10473-6
26	6	35.3	3	1	PCT-US03-11798-8

27	6	35.3	3	1	PCT-US03-14096-24	Sequence 24, App1
28	6	35.3	3	5	US-09-302-195C-1	Sequence 1, App1
29	6	35.3	3	5	US-09-716-394-11	Sequence 11, App1
30	6	35.3	3	5	US-09-853-080A-37	Sequence 37, App1
31	6	35.3	3	5	US-09-772-819-11	Sequence 11, App1
32	6	35.3	3	5	US-09-573-830-11	Sequence 11, App1
33	6	35.3	3	5	US-09-992-124B-74	Sequence 74, App1
34	6	35.3	3	5	US-09-564-045-11	Sequence 11, App1
35	6	35.3	3	5	US-09-867-847A-30	Sequence 30, App1
36	6	35.3	3	5	US-09-980-347-6	Sequence 6, App1
37	6	35.3	3	5	US-09-723-437B-11	Sequence 11, App1
38	6	35.3	3	5	US-09-939-481-81	Sequence 81, App1
39	6	35.3	3	5	US-09-898-165B-33	Sequence 33, App1
40	6	35.3	3	5	US-09-623-548A-1053	Sequence 1053, App1
41	6	35.3	3	5	US-09-623-548A-1054	Sequence 1054, App1
42	6	35.3	3	5	US-09-623-548A-1160	Sequence 1160, App1
43	6	35.3	3	5	US-09-623-548A-1170	Sequence 1170, App1
44	6	35.3	3	6	US-10-141-531-15	Sequence 15, App1
45	6	35.3	3	6	US-10-141-531-16	Sequence 16, App1

## ALIGNMENTS

RESULT 1  
US-08-472-679H-15  
Sequence 15, Application US/08472679H  
GENERAL INFORMATION:  
APPLICANT: Cochran, Mark D.  
Junker, David E.  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 267  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pamela G. Salkeld  
STREET: 2000 Galloping Hill Road  
CITY: Kenilworth  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07033  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472.679H  
FILING DATE: 07-Jun-1995  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Salkeld, Pamela G.  
REGISTRATION NUMBER: 38,607  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 298-2135  
TELEFAX: (908) 298-5388  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 15:  
US-08-472-679H-15  
Query Match 35.3%; Score 6; DB 4; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 12 D 12  
DB 1 D 1  
RESULT 2

US-08-472-679H-23  
; Sequence 23, Application US/08472679H  
; GENERAL INFORMATION:  
; APPLICANT: Cochran, Mark D.  
; Junker, David E.  
; TITLE OF INVENTION: Recombinant Swinepox Virus  
; NUMBER OF SEQUENCES: 267  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pamela G. Salkeld  
; STREET: 2000 Galloping Hill Road  
; CITY: Kenilworth  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07033  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472,679H  
; FILING DATE: 07-Jun-1995  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Salkeld, Pamela G.  
; REGISTRATION NUMBER: 38,607  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908) 298-2135  
; TELEFAX: (908) 298-5388  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:  
US-08-472-679H-23  
Query Match 35.3%; Score 6; DB 4; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 12 D 12  
DB 1 D 1  
RESULT 3  
US-08-472-679H-33  
; Sequence 33, Application US/08472679H  
; GENERAL INFORMATION:  
; APPLICANT: Cochran, Mark D.  
; Junker, David E.  
; TITLE OF INVENTION: Recombinant Swinepox Virus  
; NUMBER OF SEQUENCES: 267  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pamela G. Salkeld  
; STREET: 2000 Galloping Hill Road  
; CITY: Kenilworth  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07033  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472,679H  
; FILING DATE: 07-Jun-1995  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Salkeld, Pamela G.  
; REGISTRATION NUMBER: 38,607  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908) 298-2135  
; TELEFAX: (908) 298-5388  
; INFORMATION FOR SEQ ID NO: 33:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 33:  
US-08-472-679H-33  
Query Match 35.3%; Score 6; DB 4; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 12 D 12  
DB 1 D 1  
RESULT 5  
US-08-472-679H-15  
; Sequence 15, Application US/08472679H

REGISTRATION NUMBER: 38,607  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 298-2135  
TELEFAX: (908) 298-5388  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 33:  
US-08-472-679H-33

Query Match 35.3%; Score 6; DB 4; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 12 D 12  
DB 1 D 1

RESULT 4  
US-08-472-679H-39  
; Sequence 39, Application US/08472679H  
; GENERAL INFORMATION:  
; APPLICANT: Cochran, Mark D.  
; Junker, David E.  
; TITLE OF INVENTION: Recombinant Swinepox Virus  
; NUMBER OF SEQUENCES: 267  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pamela G. Salkeld  
; STREET: 2000 Galloping Hill Road  
; CITY: Kenilworth  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07033  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472,679H  
; FILING DATE: 07-Jun-1995  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Salkeld, Pamela G.  
; REGISTRATION NUMBER: 38,607  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908) 298-2135  
; TELEFAX: (908) 298-5388  
; INFORMATION FOR SEQ ID NO: 39:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 39:  
US-08-472-679H-39  
Query Match 35.3%; Score 6; DB 4; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 12 D 12  
DB 1 D 1  
RESULT 5  
US-08-472-679H-15  
; Sequence 15, Application US/08472679H

GENERAL INFORMATION:  
APPLICANT: Cochran, Mark D.  
Junker, David E.  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 267  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pamela G. Salkeld  
STREET: 2000 Galloping Hill Road  
CITY: Kenilworth  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07033  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,679H  
FILING DATE: 07-Jun-1995  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Salkeld, Pamela G.  
REGISTRATION NUMBER: 38,607  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 298-2135  
TELEFAX: (908) 298-5388  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 15:  
US-08-472-679H-15

Query Match 35.3%; Score 6; DB 4; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12  
DB 1 D 1

RESULT 6  
US-08-472-679H-23  
Sequence 23, Application US/08472679H  
GENERAL INFORMATION:  
APPLICANT: Cochran, Mark D.  
Junker, David E.  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 267  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pamela G. Salkeld  
STREET: 2000 Galloping Hill Road  
CITY: Kenilworth  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07033  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,679H  
FILING DATE: 07-Jun-1995  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Salkeld, Pamela G.  
REGISTRATION NUMBER: 38,607  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 298-2135  
TELEFAX: (908) 298-5388  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 33:  
US-08-472-679H-33

TELEPHONE: (908) 298-2135  
TELEFAX: (908) 298-5388  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 23:  
US-08-472-679H-23

Query Match 35.3%; Score 6; DB 4; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12  
DB 1 D 1

RESULT 7  
US-08-472-679H-33  
Sequence 33, Application US/08472679H  
GENERAL INFORMATION:  
APPLICANT: Cochran, Mark D.  
Junker, David E.  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 267  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pamela G. Salkeld  
STREET: 2000 Galloping Hill Road  
CITY: Kenilworth  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07033  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,679H  
FILING DATE: 07-Jun-1995  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Salkeld, Pamela G.  
REGISTRATION NUMBER: 38,607  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 298-2135  
TELEFAX: (908) 298-5388  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 33:  
US-08-472-679H-33

Query Match 35.3%; Score 6; DB 4; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 D 12  
DB 1 D 1

RESULT 8  
US-08-472-679H-39  
Sequence 39, Application US/08472679H  
GENERAL INFORMATION:  
APPLICANT: Cochran, Mark D.

Junker, David E.  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 267  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pamela G. Salkeld  
STREET: 2000 Galloping Hill Road  
CITY: Kenilworth  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07033  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,679H  
FILING DATE: 07-Jun-1995  
CLASSIFICATION: <UNKNOWN>  
ATTORNEY/AGENT INFORMATION:  
NAME: Salkeld, Pamela G.  
REGISTRATION NUMBER: 38,607  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 298-2135  
TELEFAX: (908) 298-5388  
INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 39:  
US-08-472-679H-39

Query Match 35.3%: Score 6; DB 4; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
DB 1 D 1

RESULT 9  
US-10-187-339-6  
; Sequence 6, Application US/10187339  
; GENERAL INFORMATION:  
; APPLICANT: Ream, Walt et al.  
; TITLE OF INVENTION: Plants Having Enhanced Gall Resistance and  
; TITLE OF INVENTION: Methods and Compositions for Producing the Same  
; FILE REFERENCE: 53629  
; CURRENT APPLICATION NUMBER: US/10/187,339  
; CURRENT FILING DATE: 2002-06-28  
; PRIOR APPLICATION NUMBER: US/09/434,837  
; PRIOR FILING DATE: 1999-11-04  
; PRIOR APPLICATION NUMBER: 60/107,185  
; PRIOR FILING DATE: 1998-11-05  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Agrobacterium tumefaciens  
US-10-187-339-6

Query Match 35.3%: Score 6; DB 6; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
DB 2 D 2

RESULT 10  
US-10-109-048-461  
; Sequence 461, Application US/10109048  
; GENERAL INFORMATION:  
; APPLICANT: COMFORT, PADMA  
; APPLICANT: KEELING, PETER L.  
; APPLICANT: RAMIREZ, NONA  
; APPLICANT: MCKEAN, ANGELA  
; APPLICANT: GAO, ZHONG  
; APPLICANT: GUAN, HANPING  
; TITLE OF INVENTION: GLUCAN CHAIN LENGTH DOMAINS  
; FILE REFERENCE: 2461-76  
; CURRENT APPLICATION NUMBER: US/10/109,048  
; CURRENT FILING DATE: 2003-03-04  
; PRIOR APPLICATION NUMBER: 60/279,720  
; PRIOR FILING DATE: 2001-03-30  
; NUMBER OF SEQ ID NOS: 1154  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 461  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: Accession No. 3832512  
US-10-109-048-461

Query Match 35.3%: Score 6; DB 6; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
DB 2 D 2

RESULT 11  
US-10-405-027-5115  
; Sequence 5115, Application US/10405027  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Human Secreted Proteins  
; FILE REFERENCE: PS806P1  
; CURRENT APPLICATION NUMBER: US/10/405,027  
; CURRENT FILING DATE: 2003-04-07  
; PRIOR APPLICATION NUMBER: 60/369,608  
; PRIOR FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: 60/376,175  
; PRIOR FILING DATE: 2002-04-30  
; NUMBER OF SEQ ID NOS: 5810  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 5115  
; LENGTH: 2  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-405-027-5115

Query Match 35.3%: Score 6; DB 6; Length 2;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
DB 1 D 1

RESULT 12  
PCT-US02-14358-15  
; Sequence 15, Application PC/TUS0214358  
; GENERAL INFORMATION:  
; APPLICANT: XENCOR  
; APPLICANT: SYNGENTA PARTICIPATIONS AG  
; APPLICANT: Briggs, Steven P.



APPLICANT: Dalmia, Bipin K.  
APPLICANT: del Val, Greg  
APPLICANT: Desjarlais, John R.  
APPLICANT: Helfetz, Peter  
APPLICANT: Luginduhl, Peter  
APPLICANT: Muchhal, Umesh  
TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity  
FILE REFERENCE: FP-71457-2-PC/RPT/RMS/RMK  
CURRENT APPLICATION NUMBER: PCT/US02/14358  
PRIOR FILING DATE: 2002-11-25  
PRIOR APPLICATION NUMBER: US 60/370,609  
PRIOR FILING DATE: 2002-04-05  
PRIOR APPLICATION NUMBER: US 60/289,029  
PRIOR FILING DATE: 2001-05-04  
NUMBER OF SEQ ID NOS: 208  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 15  
LENGTH: 3  
TYPE: PRT  
ORGANISM: Escherichia coli  
PCT-US02-14358-15

Query Match 35.3%; Score 6; DB 1; Length 3;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
|  
Db 1 D 1

RESULT 13  
PCT-US02-14358-16  
Sequence 16, Application PC/TUS0214358  
GENERAL INFORMATION:  
APPLICANT: XENCOR  
APPLICANT: SYNGENTA PARTICIPATIONS AG  
APPLICANT: Briggs, Steven P.  
APPLICANT: Dalmia, Bipin K.  
APPLICANT: del Val, Greg  
APPLICANT: Desjarlais, John R.  
APPLICANT: Helfetz, Peter  
APPLICANT: Luginduhl, Peter  
APPLICANT: Muchhal, Umesh  
TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity  
FILE REFERENCE: FP-71457-2-PC/RPT/RMS/RMK  
CURRENT APPLICATION NUMBER: PCT/US02/14358  
PRIOR FILING DATE: 2002-11-25  
PRIOR APPLICATION NUMBER: US 60/370,609  
PRIOR FILING DATE: 2002-04-05  
PRIOR APPLICATION NUMBER: US 60/289,029  
PRIOR FILING DATE: 2001-05-04  
NUMBER OF SEQ ID NOS: 208  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 16  
LENGTH: 3  
TYPE: PRT  
ORGANISM: Bacillus subtilis  
PCT-US02-14358-16

Query Match 35.3%; Score 6; DB 1; Length 3;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
|  
Db 1 D 1

RESULT 14  
PCT-US02-14358-17  
Sequence 17, Application PC/TUS0214358  
GENERAL INFORMATION:

APPLICANT: XENCOR  
APPLICANT: SYNGENTA PARTICIPATIONS AG  
APPLICANT: Briggs, Steven P.  
APPLICANT: Dalmia, Bipin K.  
APPLICANT: del Val, Greg  
APPLICANT: Desjarlais, John R.  
APPLICANT: Helfetz, Peter  
APPLICANT: Luginduhl, Peter  
APPLICANT: Muchhal, Umesh  
TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity  
FILE REFERENCE: FP-71457-2-PC/RPT/RMS/RMK  
CURRENT APPLICATION NUMBER: PCT/US02/14358  
PRIOR FILING DATE: 2002-11-25  
PRIOR APPLICATION NUMBER: US 60/370,609  
PRIOR FILING DATE: 2002-04-05  
PRIOR APPLICATION NUMBER: US 60/289,029  
PRIOR FILING DATE: 2001-05-04  
NUMBER OF SEQ ID NOS: 208  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 17  
LENGTH: 3  
TYPE: PRT  
ORGANISM: Mycobacterium leprae  
PCT-US02-14358-17

Query Match 35.3%; Score 6; DB 1; Length 3;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
|  
Db 1 D 1

RESULT 15  
PCT-US02-14358-18  
Sequence 18, Application PC/TUS0214358  
GENERAL INFORMATION:  
APPLICANT: XENCOR  
APPLICANT: SYNGENTA PARTICIPATIONS AG  
APPLICANT: Briggs, Steven P.  
APPLICANT: Dalmia, Bipin K.  
APPLICANT: del Val, Greg  
APPLICANT: Desjarlais, John R.  
APPLICANT: Helfetz, Peter  
APPLICANT: Luginduhl, Peter  
APPLICANT: Muchhal, Umesh  
TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity  
FILE REFERENCE: FP-71457-2-PC/RPT/RMS/RMK  
CURRENT APPLICATION NUMBER: PCT/US02/14358  
PRIOR FILING DATE: 2002-11-25  
PRIOR APPLICATION NUMBER: US 60/370,609  
PRIOR FILING DATE: 2002-04-05  
PRIOR APPLICATION NUMBER: US 60/289,029  
PRIOR FILING DATE: 2001-05-04  
NUMBER OF SEQ ID NOS: 208  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 18  
LENGTH: 3  
TYPE: PRT  
ORGANISM: Sarccharomyces cerevisiae  
PCT-US02-14358-18

Query Match 35.3%; Score 6; DB 1; Length 3;  
Best Local Similarity 100.0%; Pred. No. 7.6e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 D 12  
|  
Db 1 D 1

Search completed: June 5, 2003, 15:52:28

Fri Jun 6 09:14:59 2003

Job time : 35.1818 secs

us-09-150-947f-18.rapn

Page 6